

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Biocceleration Ltd.

Om protein - protein search, using SW model

Run on: March 27, 2006, 12:01:34 ; Search time 19.9701 Seconds  
(without alignment<sup>b</sup>)

554.076 Million cell updates/sec

Title: US-09-824-787B-2  
Perfect score: 597  
Sequence: 1 MSGRPQQTSVAPPBEEVERG.....ASNGETLEKITNSRPPCVIL 115

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : PIR 80.4

1: pir1:  
2: pir2:  
3: pir3:  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No. Score Query Match Length DB ID

#### Description

1 92 15.4 255 2 T19747

hypothetical prote

2 92 15.4 258 2 T19755

hypothetical prote

3 90 15.1 101 2 AD2604

conserved hypoth

4 90 15.1 101 2 D97386

hypothetical prote

5 83 13.9 96 2 DB3228

hypothetical prote

6 81 13.6 97 2 CB2257

selangprotein W-re

7 80 13.4 243 2 AV2669

probable glutaredo

8 79 13.2 395 2 BE6152

chalcone synthase

9 79 13.2 469 2 DT2722

probable MBSA prot

10 78.5 13.1 101 2 AJ3542

hypothetical prote

11 78 13.1 3670 2 T36249

CDA peptide synth

12 76 12.7 232 2 T21526

hypothetical prote

13 75.5 12.6 393 1 EG4485

phosphoribosylglyc

14 74 12.4 698 2 AD2985

oxidoreductase Atu

15 74 12.4 729 2 C98298

probable oxidore

16 74 12.4 969 2 AJ15634

probable oxidore

17 73 12.2 163 2 DB7713

probable oxidore

18 72.5 12.1 428 2 BG4301

probable oxidore

19 72.5 12.1 348 2 FT8230

alcohol dehydrogen

20 72 12.1 525 2 FT0058

gamma-glutamyltran

21 72 12.1 938 2 AJ13417

glutamate-ammonia

22 72 12.1 987 2 148953

eph-related recept

23 72 12.1 1507 2 D97106

large chain of NAD

24 72 12.0 347 2 G33171

hypothetical prote

25 71.5 11.9 413 2 T03270

probable histidino

26 71 11.9 448 1 C69594

adenosylmethionine

27 70.5 11.8 282 2 F86577

elongation factor

28 70.5 11.8 467 2 C72601

probable tryptopho

#### ALIGNMENTS

#### RESULT 1

T19747

hypothetical protein C35C5.3a - Caenorhabditis elegans

C;Species: Caenorhabditis elegans #Sequence\_revision 15-Oct-1999 #text\_change 18-Feb-2000

C;Accession: T19747

R;White, S.

Submitted to the EMBL Data Library, August 1996

A;Reference number: Z19173

A;Accession: T19755

A;Status: preliminary; translated from GB/EMBL/DBDJ

A;Molecule type: DNA

A;Residues: 1-258 <WIL>

A;Cross-references: UNIPARC:UPI000164318; EMBL:Z78417; PIDN:CA01692.1; GSPDB:GN00028; A;Experimental source: clone C35C5

chalcone isomerase  
hypothetical prote  
probable 3-hydroxy  
outer membrane pro  
tein F9C16.9 [1  
hypothetical prote  
involved in spore  
molybdenum cofacto  
phosphoribosylglyc  
PRM oxidoreductase  
proline dehydrogen  
hypothetical prote  
hypothetical prote  
molybdopterin conv  
methyldopterin conv  
hypothetical prote

30	70	11.7	241	2	T08006
31	70	11.7	489	2	C69433
32	70	11.7	504	2	T36703
33	70	11.7	673	2	S60142
34	70	11.7	946	2	D98503
35	69.5	11.6	167	2	G98828
36	69.5	11.6	411	2	H3788
37	69	11.6	159	2	E75151
38	69	11.6	419	2	D70632
39	69	11.6	430	2	P87472
40	69	11.6	1310	1	I53597
41	68.5	11.5	207	2	G8971
42	68.5	11.5	208	2	T12985
43	68.5	11.5	239	2	C65220
44	68	11.4	139	2	H72320
45	68	11.4	183	2	F84636



A;Map position: 1

Query Match 13.6%; Score 81; DB 2; Length 97;  
 Best Local Similarity 29.3%; Pred. No. 0.78;  
 Matches 22; Conservative 14; Mismatches 33; Indels 6; Gaps 3;

Qy 24 RIIVVYCEPGGF--EATYL--ELASAVKEGYPPGIESERUGGTAFAEIEINGQVFSKLE 79  
 S QEIVTYCROCNWMLRSAWLSQLQELHMTSEELEYVALHPDTG--GRFEIFCNGVQIWERKO 62

Qy 80 NGGFIVYEKOLIEAIR 94

Db 63 EGGFPPEAKVILQVR 77

RESULT 7

probable glutaredoxin-like protein APE0775 - Aeropyrum pernix (strain K1)

C;Species: Aeropyrum pernix

C;Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 09-Jul-2004

C;Accession: A72669

R;Kawarabayasi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takah

ava, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudo, Y.; Yamazaki, J.; K

DNA Res. 6, 83-101, 1999

A;Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyr

A;Reference number: A72450; MUID:99310339; PMID:10382966

A;Accession: A72669

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-243 &lt;KAW&gt;

A;Cross-references: UNIPROT:Q9YD24; UNIPARC:UPI000005DCA5; DDBJ:AP000060; NID:95104188;

A;Experimental source: strain K1

A;Gene: APE0775

Query Match 13.4%; Score 80; DB 2; Length 243;

Best Local Similarity 34.9%; Pred. No. 2.8%;  
 Matches 22; Conservative 9; Mismatches 26; Indels 6; Gaps 2;

Qy 42 LASAVKQPGTIESRUGGTGAFETEINGQVFSKLENGGFPYKEKDLEATIRASNGTET 101

Db 176 LSEAV-EATENPDIADPKVMSVSIATNGYLVNV----GVYEEEDFLDVYVSAEERL 229

RESULT 8

Chalcone synthase homolog RT123.4 - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 05-Oct-2004

C;Accession: E86152

R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, C.;Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;ansen, N.F.; Hughes, B.; Ruizar, L.; Nature, 408, 816-820, 2000

C;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.C.A.; Di, J.H.; Di, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marzilli, R.;Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.; Salzberg, S.L.; Schwartz, J.R.; Shin, P.; Southwick, A.M.; Sun, H.; Tallon, A.;Tucker, M.; Wu, D.; Yu, J.; Fraser, C.M.; Venter, J.C.; Davis, R.W.; Takemoto, K.;A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A;Reference number: A86141; MUID:21016719; PMID:1130712

A;Accession: E86152

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-395 &lt;STO&gt;

A;Cross-references: UNIPARC:UPI00000481A3; GB:AE005172; NID:92317904; PIDN: AAC24368.1;

C;Genetics: C;Map position: 1

C;Superfamily: type III polyketide synthase

Query Match 13.2%; Score 79; DB 2; Length 395;  
 Best Local Similarity 24.1%; Pred. No. 6;  
 Matches 32; Conservative 15; Mismatches 36; Indels 50; Gaps 5;

Qy 13 PPEVEP----GIVGIVVYECSPCGFATYLLILASAVKEQYQPGIE--IESRL-- 59  
 Db 203 PPNKARPVDLVGALFGDGAAVIGADFRECSPFMELHYAVQOFLPDTQVIEGRLT 267

Qy 60 -----GTTGATIEIN-----GOLVFSKLEIN-----GOLVFSKLEIN 83

Db 268 EGTFPKLGDLPKIENTEIFCCKLMKAGDSMWFNUFWVHPGGPAPLRLT-KL 326

Db 327 KLEKEKUSSRRA 339

RESULT 9

probable MRSA protein APE0317 - Aeropyrum pernix (strain K1)

C;Species: Aeropyrum pernix

C;Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 09-Jul-2004

C;Accession: D72722

R;Kawarabayasi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takah

awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudo, Y.; Yamazaki, J.; K

DNA Res. 6, 83-101, 1999

A;Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyr

A;Reference number: A72450; MUID:99310339; PMID:10382966

A;Accession: D72722

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-469 &lt;KAW&gt;

A;Cross-references: UNIPROT:Q9YFC4; UNIPARC:UPI000005DCA5; DDBJ:AP000059; NID:95103911;

A;Experimental source: strain K1

A;Gene: APE0317

C;Superfamily: phosphomannomutase

Query Match 13.2%; Score 79; DB 2; Length 469;

Best Local Similarity 27.6%; Pred. No. 7.3; Mismatches 35; Indels 40; Gaps 7;

Matches 35; Conservative 16; Mismatches 36; Indels 40; Gaps 7;

Qy 1 MSGPQGTTSVAPPVEVPGSSVRIWV-YCSPCGFATYLE----LASAVKQYQFIE 54  
 Db 273 LAGFVAVGVLAKGPKLYTGHSSVIVEDYLPRPGTVWTVGAPITARALITQ---- 327

Qy 55 IESRLGGTGAFFETEINGQVFSKLENGGFP----PYKEKD----LITEATIRASNGTET 103

Db 328 ----GGTAFF-----ENGGYMPHOTVRDGGWKAALLLAMLRAT-GESLS 369

Qy 104 KITNSRP 110  
 Db 370 SLDRLP 376

RESULT 10

hypothetical protein BMER10262 [imported] - Brucella melitensis (strain 16M)

C;Species: Brucella melitensis

C;Date: 01-Feb-2002 #sequence\_revision 01-Feb-2002 #text\_change 09-Jul-2004

C;Accession: AD3542

R;Del Vecchio, V.G.; Kapatral, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova, N.; Mazur, M.; Gotsman, R.; Seikov, E.; Elzer, P.H.; Regius, S.; O'Callaghan, D.; Lees, J.; Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002

A;Title: The genome sequence of the facultative intracellular pathogen Brucella melitensis

A;Reference number: AD3252; PMID:11756688

A;Accession: AD3542

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-101 &lt;KAW&gt;

A;Cross-references: UNIPROT:Q9YDB6; UNIPROT:Q8FV19; UNIPARC:UPI00000583P4; GB:AE008918;

C;Genetics: C;Map position: 1

C;Superfamily: type III polyketide synthase

C;Genetics: A;Gene: EMBI0262  
A;Map position: II

Query Match 13.1%; Score 78.5; DB 2; Length 101;  
Best Local Similarity 26.9%; Pred. No. 1.4;  
Matches 21; Conservative, 16; Mismatches 34; Indels 7; Gaps 2;

Query Match 12.7%; Score 76; DB 2; Length 232;  
Best Local Similarity 15.1%; Pred. No. 6.4;  
Matches 21; Conservative, 19; Mismatches 31; Indels 68; Gaps 1;

Query Match 24 RIVVEYCEPCGCFEATYLELASAVKEQYQPG--IEESRLGGTGAETEI----NGQLVFS 76  
Db 6 RISITVCTQCNWLRRAWMAQDIELLQFGQDIALEVALRPGFGVFEIRVQMDGSBELINE 65

Query Match 56 INLTIVFSVSCGKIQAFNQFYEFAKEKIPGLVIEGGNFSSPDFWKGLCLAQIVGAKIGIA 145  
Db 146 IVITGSNPFPFYIGFQYQIQLQTAHYNRFSYSSYLVMIGNLFESTLSITCAFEIFLGDKOI 205

Query Match 75 FSKELENGGPPYKEDILEAT 93  
Db 205 WSKISKERVPTQEEFLNLI 224

RESULT 11

T36249 CDA peptide synthetase II SCB63.02c [imported] - Streptomyces coelicolor  
C;Species: Streptomyces coelicolor  
C;Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 12-Jul-2004  
C;Accession: T36249  
R;Saunders, D.C.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.  
submitted to the EMBL Data Library, March 1999  
A;Reference number: Z21602  
A;Status: preliminary; translated from GB/EMBL/DDBJ  
A;Molecule type: DNA  
A;Residues: 1-3670 <SAU>  
A;Cross-references: UNIPROT:Q9Z4X5; UNIPARC:UPI00000DAF1; EMBL:AL035640; PIDN: CAB38517  
A;Experimental source: Strain A3(2)  
C;Genetics:

Query Match 13.1%; Score 78; DB 2; Length 3670;  
Best Local Similarity 31.7%; Pred. No. 97;  
Matches 26; Conservative, 14; Mismatches 30; Indels 12; Gaps 3;

Query Match 12.6%; Score 75.5; DB 1; Length 393;  
Best Local Similarity 22.5%; Pred. No. 13;  
Matches 31; Conservative, 23; Mismatches 37; Indels 47; Gaps 7;

Query Match 1 1 MSBRPGQTSVAPSPPEEV-----GSGYRIVWE-----Y 29  
Db 161 MSSSGKGGSQSVRSBEDIEKAWKTAKEGARGRIGRNIVVIEBFINFDYETIILTARTAEGTRP 220

RESULT 12

T21526 hypothetical protein F28H7.4 - Caenorhabditis elegans  
C;Species: Caenorhabditis elegans  
C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
C;Accession: T21526  
R;Barks, M.  
Submitted to the EMBL Data Library, May 1996  
A;Reference number: Z19435  
A;Accession: T21526  
A;Status: preliminary; translated from GB/EMBL/DDBJ  
A;Molecule type: DNA  
A;Residues: 1-232 <WIL>  
A;Cross-references: UNIPROT:Q19892; UNIPARC:UPI000135A34; EMBL:Z72508; PIDN:CAA96637.1;  
C;Experimental source: clone F28H7  
C;Genetics:

RESULT 14

AD2985 oxidoreductase Atu3485 [imported] - Agrobacterium tumefaciens (strain C58, dupont)  
C;Species: Agrobacterium tumefaciens (strain C58, dupont)  
C;Date: 11-Jan-2002 #sequence\_revision 11-Jan-2002 #text\_change 09-Jul-2004  
C;Accession: AD2985

A;Gene: CESPF28H7.4  
A;Map position: 5  
A;Introns: 23/1; 134/3

Query Match 23 VRIVVEYCEPCGCFEATYLELASAVKEQYQPG--IEESRLGGTGAETEI----NGQLVFS 55  
Db 86 INLTIVFSVSCGKIQAFNQFYEFAKEKIPGLVIEGGNFSSPDFWKGLCLAQIVGAKIGIA 145

Query Match 55 IVITGSNPFPFYIGFQYQIQLQTAHYNRFSYSSYLVMIGNLFESTLSITCAFEIFLGDKOI 205

Query Match 75 FSKELENGGPPYKEDILEAT 93  
Db 205 WSKISKERVPTQEEFLNLI 224

RESULT 13

E64485 phosphoribosylglycinamide formyltransferase 2 (EC 2.1.1.2.-) MJ1486 [similarity] - Methanococcus jannaschii  
C;Species: Methanococcus jannaschii  
C;Accession: E64485  
R;Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake, J.; Reich, C.I.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek, A.; Rosen, J.D.; Sadow, P.W.; Hama, M.; Cotton, K.M.; Hurst, M.A.  
Science 273, 1058-1073, 1996  
A;Authors: Kline, B.; Borodovsky, M.; Kleink, H.P.; Fraser, C.M.; Smith, H.O.; Woese, C.  
A;Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii  
A;Accession: E64485  
A;Status: nucleic acid sequence not shown; translation not shown  
A;Molecule type: DNA  
A;Residues: 1-393 <BUL>  
A;Reference number: A62300; MURB:96337999; PMID:8688087  
C;Genetics:

Query Match 1 1 MSBRPGQTSVAPSPPEEV-----GSGYRIVWE-----Y 29  
Db 161 MSSSGKGGSQSVRSBEDIEKAWKTAKEGARGRIGRNIVVIEBFINFDYETIILTARTAEGTRP 220

Query Match 30 CEPGQ--PEATLE-----LASAVKEQYQPG--IEESRLGGTGAETEI----NGQLVFS 75  
Db 221 CEPIGHVQDGDHESWQPHNMSAELKEQDIAKKVTDALGGYGYIGFVLFVKGDEVIF 280

Query Match 76 SKLENGGPPYKEDILEAT 93  
Db 281 SEVSPR--PDHDTGMVMI 296

R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monkem, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, L.; Erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McClellan, R.; Karp, P.; Romero, P.; Zhang, S.; Science 294, 2317-2323, 2001  
 A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm, R.; Author: E.W.;  
 A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.  
 A;Reference number: AB2577; MVID:21608550; PMID:1743193  
 A;Accession: AD2985  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residue: 1-698 <KUR>  
 A;Cross-references: UNIPROT:Q8UAB9; UNIPARC:UPI000016480F; GB:AE008699; PIDN:AAI44298.1;  
 A;Experimental source: strain C58 (Dupont)  
 A;Gene: Atu3485  
 A;Map position: linear chromosome

Query Match 12.4%; Score 74; DB 2; Length 698;  
 Best Local Similarity 22.9%; Pred. No. 35; Matches 25; Conservat. 20; Mismatches 38; Indels 26; Gaps 3;

QY 16 EVEPGSGVRLVVEYEP-----GREATVLELASAVKEQYPGIETESRLGG---- 61  
 Db 53 BIDELTGEKURVMMWKGSGGDVGTIKLDGFTALYDKEALKGYKGVEDEDRMVGFPH 112  
 QY 62 -----TCAFEETINGOLVFSKLENGGPPYEKOLIEARRASNGETL 102  
 Db 113 CTFNLNPRASAIDTPLHGFVPPFDHVH---MHPDAITAASKSKEL 157

RESULT 15

C98298 probable oxidoreductase [imported] - Agrobacterium tumefaciens (strain C58, Cereon)  
 C;Species: Agrobacterium tumefaciens  
 C;Date: 22-Oct-2001 #sequence\_revision 22-Oct-2001 #text\_change 09-Jul-2004  
 C;Accession: C98298  
 R;Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Ourolo, B.; Goldman, A.; Liu, F.; William, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.; Science 294, 2323-2328, 2001  
 A;Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum  
 A;Reference number: A97359; MVID:2160551; PMID:11743194  
 A;Accession: C98298  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-729 <KUR>  
 A;Cross-references: UNIPROT:Q8UAB9; UNIPARC:UPI000002209; GB:AE007870; PIDN:AAK89909.1;  
 A;Gene: AGR\_L\_2690  
 A;Map Position: linear chromosome

Query Match 12.4%; Score 74; DB 2; Length 729;  
 Best Local Similarity 22.9%; Pred. No. 37; Matches 25; Conservat. 20; Mismatches 38; Indels 26; Gaps 3;

QY 16 EVEPGSGVRLVVEYEP-----GREATVLELASAVKEQYPGIETESRLGG---- 61  
 Db 84 EIDPLTGEKURVMMWKGSGGDVGTIKLDGFTALYDKEALKGYKGVEDEDRMVGFPH 143  
 QY 62 -----TCAFEETINGOLVFSKLENGGPPYEKOLIEARRASNGETL 102  
 Db 144 CTFNLNPRASAIDTPLHGFVPPFDHVH---MHPDAITAASKSKEL 188

Search completed: March 27, 2006, 12:08:35  
 Job time : 22.9701 secs

GenCore version 5.1.7  
Copyright: (c) 1993 - 2006 Biocceleration Ltd.

Om protein - protein search, using sw model

Run on: March 27, 2006, 12:00:46 ; Search time 135.659 Seconds  
(without alignments)  
598.087 Million cell updates/sec

Title: US-09-824-787B-2  
Perfect score: 597

Sequence: 1 MSGPGQI3VAPPPEEVPG.....ASNGTELKITNSRPPCVIL 115

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched:

2166443 Begg, 705528306 residueB

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Maximum Match 0%

Listing first 45 summaries

Database :

UniProt 05.80:\*

1: uniprot\_sprot:\*

2: uniprot\_trembl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query	Match Length	DB ID	Description
1	597	100.0	115	2	Q9BRT3_HUMAN
2	537	89.9	115	2	Q9CQ86_MOUSE
3	435.5	72.9	126	2	Q5ZH7H_CHICK
4	273	45.7	95	2	Q82F5P_BRARE
5	221.5	37.1	94	2	Q8QG8P_BRARE
6	218	36.5	89	2	Q7P504_TETNG
7	148.5	24.9	95	2	Q9VRAO_DROME
8	137.5	23.0	88	2	Q8I6T4_CHLRE
9	137	22.9	80	2	Q4B6M6_SULAC
10	132.5	22.2	93	2	Q7QFL8_ANOGA
11	119	23.2	2	Q8Z227_ORYSA	
12	114.5	19.2	24	1	HS26_HENGL
13	112.5	18.8	86	1	SE911_HUMAN
14	112.5	18.8	86	1	SE911_MACMU
15	112.5	18.8	86	1	SE911_PIG
16	112.5	18.8	86	2	Q8Q2G9_BRARE
17	112.5	18.8	87	2	Q5NVT2_PONPY
18	111	18.5	209	2	Q67ZS9_BRATH
19	110.5	18.5	86	1	SE911_HUMAN
20	110.5	18.5	86	1	SE911_MACMU
21	109	18.3	329	2	Q80112_MOUSE
22	107.5	18.0	87	1	SE911_MOUSE
23	107.5	18.0	87	1	SE911_RAT
24	104	17.4	257	17	Q65117_ORYSA
25	103.5	17.3	228	2	Q81B55_ARATH
26	103	16.6	99	2	Q8D8L8_VIBVU
27	99	16.6	94	2	Q7M811_VIBVU
28	98.5	16.5	86	1	SE911_SHEP
29	94	15.7	92	2	Q87RH7_VIBVU
30	94	10.3	157	103	Q60CA7_METCPA
31	93	15.6	93	2	Q92SU7_RHIME

Alignments

RESULT 1	Q9BRT3_HUMAN	Q9BRT2_HUMAN PRELIMINARY;	PRT;	115 AA.
ID	Q9BRT3;			
AC				
DT	01-JUN-2001 (T=EMBL; 17, last sequence update)			
DT	10-MAY-2005 (T=EMBL; 30, last annotation update)			
DE	Chromosome 17, open reading frame 37 (XTP4) (C35 protein).			
GN	Name=C17orf37; Synonyms=XTP4;			
OS	Homo sapiens (Human).			
OC	Bukaryota; Metzozoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Buteraria; Burchontoglires; Primates; Catarrhini; Hominidae; Homo.			
OX				
RN	[1]			
RP	NUCLEOTIDE SEQUENCE.			
RC	TISSUE=Brain;			
RA	MEIDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;			
RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,			
RA	Klaunser R.D., Collins F.S., Scheiner C.M., Schuler G.D.,			
RA	Altschul S.F., Zeeberg B., Buetow K.H., Schaeffer C.F., Bht N.K.,			
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,			
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,			
RA	Stapleton M., Soares M.B., Bernaldo M.F., Casavant T.L., Schatz T.E.,			
RA	Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,			
RA	Raha S.S., Loquaiello N.A., Peters G.J., Abramson R.D., Mullahy S.J.,			
RA	Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunnarsson P.H.,			
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,			
RA	Villalon D.K., Muzny D.M., Soedergren E.J., Lu X., Gibbs R.A.,			
RA	Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,			
RA	Whiting M., Madan A., Young A.C., Shevchenko K.Y., Bouffard G.G.,			
RA	Blakesley R.W., Tochman J.W., Green E.D., Dickson M.C.,			
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,			
RA	Butterfield Y.S.N., Krzywinski M.I., Salska U., Smailus D.E.,			
RA	Schnerich A., Schein J.E., Jones S.M., Marras M.A.,			
RA	RT "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences";			
RT	and mouse genomic sequences".			
RT	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).			
RN	[2]			
RP	NUCLEOTIDE SEQUENCE.			
RC	TISSUE=Brain.			
RA	Director MGC Project;			
RL	Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.			
RL	[3]			
RP	NUCLEOTIDE SEQUENCE.			
RA	NUCLEOTIDE SEQUENCE.			
RA	RA Liu Y., Cheng J., Lu Y., Wang G., Zhang L., Chen J., Li L.,			
RA	Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.			
RL	[4]			
RA	NUCLEOTIDE SEQUENCE.			
RA	RA Evans E., Henm A.D., Luhowskyj S., Paris M.J., Borrello M.A.,			
RA	Smith B.S., Sahasrabudhe D.M., Zauderer M.,			
RL	Submitted (DEC-2003) to the EMBL/GenBank/DDBJ databases.			
DR	DR EMBL; BC006005; AAB06001; -; mRNA.			
DR	EMBL; AF490253; AAO85461.1; -; mRNA.			
DR	Q92SU7_RHIME			



RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=C57BL/6J; TISSUE=Aorta and vein;  
 RA Adachi J., Mizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,  
 RA Fukuda S., Furuno M., Hangaki T., Hara A., Hashizume W.,  
 RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoaka T., Hirozane T.,  
 RA Hori F., Imotani K., Ishii Y., Itoh M., Kasukawa T.,  
 RA Katoh H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,  
 RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,  
 RA Nishi K., Nomura R., Numazaki R., Ohno M., Ohnato N., Okazaki Y.,  
 RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,  
 RA Sabaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,  
 RA Tagawa A., Takahashi T., Takaku-Akaiwa S., Takeda Y., Tanaka T.,  
 RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;  
 RL Submitted (JUL-2001) to the EMBL/GenBank/DDBJ databases.  
 RN [8]

RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=FVB/N; TISSUE=Liver;  
 RD MEDLINE=22188257; PubMed=2477932; DOI=10.1073/pnas.242603899;  
 RA strauberg R.L., Feingold B.A., Grouse L.H., Derge J.G.,  
 RA Klaushier R.D., Collins F.S., Wegner L., Schmenner C.M., Schuler G.D.,  
 RA Aitachul S., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heilek P.,  
 RA Datchenko N., Marusina K., Farmer A.R., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Schatz T.E.,  
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Rosa S.S., Loquellano N.A., Peters G.J.J., Abramson R.D., Mullany S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Heitton C., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Buffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalius D.E.,  
 RA Sturner A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 and mouse cDNA sequences";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
 RN [9]

RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=FVB/N; TISSUE=Liver;  
 RA Strauberg R.;  
 RL Submitted (JAN-2002) to the EMBL/GenBank/DDBJ databases.  
 DR EMBL; AK00922; BAB2486.1; -; mRNA.  
 DR EMBL; AK00929; BAB2480.1; -; mRNA.  
 DR EMBL; AK01779; BAB2561.1; -; mRNA.  
 DR EMBL; AK01134; BAC0901.1; -; mRNA.  
 DR EMBL; BC021589; AAH21589.1; -; mRNA.  
 DR Ensemble; ENSMUSG0000002590; Mus musculus.  
 DR MGI; MGI:1313678; 181046019RICK.  
 DR InterPro; IPR011893; CXXU\_se1WTH; 1.  
 DR TIGRFAMS; TIGR02174; CXXU\_se1WTH; 1.  
 KW Hypothetical protein.  
 SQ SEQUENCE 115 AA; 12295 MW; B36A0340DFEB737A CRC64;

RESULT 4  
 Query Match 89.9%; Score 537; DB 2; Length 115;  
 Best Local Similarity 89.6%; Pred. No. 6\_1e-45;  
 Matches 103; Conservative 4; Mismatches 8; Indels 0; Gaps 0;  
 RN

QY 2 SGBPGQTSAVAPPPEVPGSG-----VRIIVEYCPGCGTATVIELASAVKEQ 49  
 Db 4 SGNGAAVG--TBSSEAGGDDGFGSSDGGSSRRVHIMVECEPCPFGATEGATEELASAVRE 60  
 DR 01-JUN-2003 (TREMBrel. 24; Last sequence update)  
 DT 01-MAR-2004 (TREMBrel. 26; Last annotation update)  
 DB Selenoprotein W2a.  
 GN Name=sepW2a;  
 OS Brachyota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
 OC Cyprinidae; Danio.  
 OC NCBI\_TaxID=7955;  
 RN [1]

RP NUCLEOTIDE SEQUENCE.  
 RD MEDLINE=22796651; PubMed=12915322; DOI=10.1016/S1567-133X(03)00054-1;  
 RA Thisse C., Degrave A., Kryukov G.V., Gladyshev V.N.,  
 RA Obrecht-Pflumio S., Kroil A., Thisse B., Lescure A.;  
 RT "Spatial and temporal expression patterns of selenoprotein genes  
 during embryogenesis in zebrafish.";  
 Gene Expr. Patterns 3:525-532 (2003).  
 DR EMBL; AY221261; AA65270.1; -; mRNA.  
 DR Ensemble; ENSDARG0000021296; Danio rerio.  
 DR InterPro; IPR011893; CXXU\_se1WTH.  
 DR TIGRFAMS; TIGR02174; CXXU\_se1WTH; 1.  
 KW Selenum; Selenocysteine.  
 FT SE\_CVS 13 13  
 SQ SEQUENCE 95 AA; 10572 MW; 05A25B769DDFDDB0 CRC64;

Query Match 45.7%; Score 273; DB 2; Length 95;

RESULT 3  
 QZINH? CHICK  
 ID QZINH? CHICK PRELIMINARY; PRT; 126 AA.  
 AC QZINH?  
 DT 25-OCT-2004 (TREMBrel. 28; Last sequence update)  
 DT 25-OCT-2004 (TREMBrel. 28; Last annotation update)  
 DR Hypothetical protein.  
 GN ORFNAMES=RCOMB04\_2622;  
 OS Gallus gallus (Chicken).  
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
 OC Gallus.  
 OC Gallus.  
 RN [1]

RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=CB; TISSUE=Bursa;  
 RA Caldwell R.B., Kierzek A.M., Arakawa H., Bezzubov Y., Zaim J.,  
 RA Fiedler P., Kutter S., Blagodatski A., Kostovska D., Kotter M.,  
 RA Plachy J., Carninci P., Buerstedde J.M.;  
 RT "Full-length cDNAs from chicken bursal lymphocytes to facilitate  
 genefunction analysis.";  
 RL Genome Biol. 6:R6 (2005).  
 DR EMBL; AJ720807; CAG32466.1; -; mRNA.  
 DR InterPro; IPR011893; CXXU\_se1WTH.  
 DR TIGRFAMS; TIGR02174; CXXU\_se1WTH; 1.  
 KW Hypothetical protein.  
 SQ SEQUENCE 126 AA; 13438 MW; 12B8498FD40DBB6D CRC64;

Query Match 72.9%; Score 435; DB 2; Length 126;  
 Best Local Similarity 72.2%; Pred. No. 6.9e-35; Mismatches 16; Indels 15; Gaps 2;  
 Matches 91; Conservative 4; Mismatches 16;

QY 50 YPGIETEESRLGGTGAETBIEINGOLVTSKLEENGFFPVEKDEJTEAIRRASNGTLEKITNSR 109  
 Db 61 YPDIEIESRLGGTGAETEINGOLVTSKLEENGFFPVEKDEJTEAIRRASNGTLEKITNSR 120

QY 110 PPCVIL 115  
 Db 121 PPCTIL 126



RA Foaler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glaser K.,  
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Hostin N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Houston D., Houston K.A., Howland T.J., Wei M.-H., Ibeagwam C.,  
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Iabko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Milshina N.V., Mobarie C., Morel J., Moskrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,  
 RA Palazzo M., Pittman G.S., Pan S., Pollard J., Purif V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Schueler F., Shen H.,  
 RA Shue B.C., Siden-Kiamas I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirska R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wasserman D.A., Weinstock J.M., Wessendorf J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Yee J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibb R.A., Myers E.W., Rubin G.M., Venter J.C.,  
 RT "The genome sequence of *Drosophila melanogaster*.",  
 RL *Science* 287:2185-2195 (2000).  
 RN [2]  
 RP  
 RX MEDLINE=22426065; PubMed=12537568;  
 RA Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,  
 RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,  
 RA George R.A., Hobkins R.A., Laverty T., Munny D.M., Nelson C.R.,  
 RA Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,  
 RA Svirska R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,  
 RA Weinstock G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.,  
 RT "Finishing a whole-genome shotgun: release 3 of the *Drosophila*  
 melanogaster euchromatic genome sequence.",  
 RL *Genome Biol.* 3:RESEARCH0079 (2002).  
 RN [3]  
 RP  
 RX NUCLEOTIDE SEQUENCE.  
 RA MEDLINE=22426070; PubMed=12537573;  
 RA Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J.W., Svirska R.,  
 RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,  
 RA Ashburner M., Celniker S.E.,  
 RT "The transposable elements of the *Drosophila melanogaster* euchromatin: a genomics perspective.",  
 RL *Genome Biol.* 3:RESEARCH0084.1-RESEARCH0084.20 (2002).  
 RN [4]  
 RP NUCLEOTIDE SEQUENCE.  
 RX MEDLINE=22426069; PubMed=12537572;  
 RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,  
 RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,  
 RA Smith C.D., Tupy J.L., Whitfield E.J., Barakatoglu L., Berman B.P.,  
 RA Babbencourt D.R., Celniker S.E., de Grey A.D.N.J., Drysdale R.A.,  
 RA Harris N.L., Richter J.J., Russo S., Schroeder A.J., Shu S.Q.,  
 RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,  
 RA Lewis S.E.,  
 RT "Annotation of the *Drosophila melanogaster* euchromatic genome: a systematic review.",  
 RL *Genome Biol.* 3:RESEARCH0083.1-RESEARCH0083.22 (2002).  
 RN [5]  
 RP NUCLEOTIDE SEQUENCE.  
 RG Berkeley Drosophila Genome Project;  
 RA Celniker S., Carlson J., Wan K., Pfeiffer B., Frise E., George R.,  
 RA Hoskins R., Stapleton M., Pacleb J., Park S., Svirska R., Smith B.,  
 RT "Drosophila melanogaster release 4 sequence.",  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.  
 RN [6]  
 RP NUCLEOTIDE SEQUENCE.  
 RG FlyBase;  
 RL Submitted (MAR-2005) to the EMBL/GenBank/DDBJ databases.  
 CC -1- INTERACTION:  
 CC Q8T4D9;CG15056; NDEExp=1; IntraAct=BI-82672, EBI-135162;  
 CC EMBL; AEO0571; AAF5093.1; -, Genomic\_DNA;  
 DR Interact; Q9VRA0; -;

DR Ensembl; CG15456; *Drosophila melanogaster*.  
 DR FLYBase; Flggn0040650; CG15456.  
 DR GO; GO:005739; C:mitochondrion; IEA.  
 DR GO; GO:0005489; F:electron transporter activity; IEA.  
 DR GO; GO:0005506; F:iron ion binding; IEA.  
 DR GO; GO:0006118; P:electron transport; IEA.  
 DR GO; GO:0006810; P:transport; IEA.  
 DR Interpro; IPR01893; CXXU selfWH.  
 DR Interpro; IPR001041; Ferredoxin.  
 DR TIGRFAMS; TIGR03174; CXXU selfWH; 1  
 SQ SEQUENCE; 95 AA; 10506 MW; A4564893BF263FDS CRC64;  
 DR Query Match Similarity 24.9%; Score 148.5; DB 2; Length 95;  
 DR Best Local Similarity 34.0%; Pred. No. 9.6e-07;  
 DR Matches 32; Conservative 17; Mismatches 44; Indels 1; Gaps 1;  
 DR 01-MAR-2003 (T2EMBLrel. 23; Last sequence update)  
 DR 01-MAR-2003 (T2EMBLrel. 23; Last annotation update)  
 DE Selenoprotein SelW.  
 OS Chlamydomonas reinhardtii.  
 OC Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae;  
 OC Chlamydomonadales; Chlamydomonaceae; Chlamydomonas.  
 OX NCBI\_TaxID=3055;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RA Novoselov S.V., Rao M., Onoshko N.V., Zhi H., Kryukov G.V., Xiang Y.,  
 RA Weeks D.P., Hartfield D.L., Gladyshev V.N.,  
 RL Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.  
 DR EMBL; AF494050; AN32901.1; -, mRNA.  
 DR Interpro; IPR011893; CXXU selfWH.  
 DR TIGRFAMS; TIGR02174; CXXU\_selfWH; 1.  
 KW Selenium; Selenocysteine.  
 FT SB\_CYS 14 14  
 SO SEQUENCE 88 AA; 9690 MW; DR27CBB4790E128 CRC64;  
 DR Query Match 23.0%; Score 137.5; DB 2; Length 88;  
 DR Best Local Similarity 37.2%; Pred. No. 1.1e-05;  
 DR Matches 32; Conservative 17; Mismatches 32; Indels 5; Gaps 2;  
 DR Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,  
 DR Lewis S.E.,  
 RT "Annotation of the *Drosophila melanogaster* euchromatic genome: a systematic review.",  
 RL *Genome Biol.* 3:RESEARCH0083.1-RESEARCH0083.22 (2002).  
 RN [5]  
 RP NUCLEOTIDE SEQUENCE.  
 RG Berkeley Drosophila Genome Project;  
 RA Celniker S., Carlson J., Wan K., Pfeiffer B., Frise E., George R.,  
 RA Hoskins R., Stapleton M., Pacleb J., Park S., Svirska R., Smith B.,  
 RT "Drosophila melanogaster release 4 sequence.",  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.  
 RN [6]  
 RP NUCLEOTIDE SEQUENCE.  
 RG FlyBase;  
 RL Submitted (MAR-2005) to the EMBL/GenBank/DDBJ databases.  
 CC -1- INTERACTION:  
 CC Q8T4D9;CG15056; NDEExp=1; IntraAct=BI-82672, EBI-135162;  
 CC EMBL; AEO0571; AAF5093.1; -, Genomic\_DNA;  
 DR Interact; Q9VRA0; -;

RESULT 9  
 ID 04JBM6\_SULAC  
 ID 04JBM6\_SULAC PRELIMINARY; PRT; 80 AA.  
 AC 04JBM6\_SULAC  
 DT 13-SEP-2005 (T2EMBLrel. 31; Last sequence update)  
 DT 13-SEP-2005 (T2EMBLrel. 31; Last annotation update)  
 DE Hypothetical protein.  
 GN OrderdLocusName=Scsi0387;  
 OS Sulfolobus acidocaldarius  
 OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;







GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM protein - protein search, using sw model

Run on: March 27, 2006, 12:07:54 ; Search time 29.9551 Seconds  
(without alignments)

Scoring table: BLOSUM62

Perfect score: 597  
Sequence: 1 MSGEPGQTSVAPPPEEVPG.....ASNGETLKKITNSRPPCVIL 115

Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

Issued Patents AA:\*

- 1: /cgn2\_6/ptodata/1/iaa/5-COMB.pep:\*
- 2: /cgn2\_6/ptodata/1/iaa/H-COMB.pep:\*
- 4: /cgn2\_6/ptodata/1/iaa/PCTUS-COMB.pep:\*
- 5: /cgn2\_6/ptodata/1/iaa/RE-COMB.pep:\*
- 6: /cgn2\_6/ptodata/1/iaa/backfiled.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	403	67.5	90	2 US-09-513-999C-7234 Sequence 7234, AP
2	96.5	16.2	88	1 US-09-127-289-3 Sequence 3, Appli
3	96.5	16.2	88	1 US-09-333-448-3 Sequence 3, Appli
4	96.5	16.2	93	2 US-09-333-448-3 Sequence 1, Appli
5	96.5	16.2	93	1 US-09-127-289-1 Sequence 1, Appli
6	96.5	16.2	93	2 US-09-333-448-1 Sequence 1, Appli
7	87	14.6	216	2 US-09-270-767-46297 Sequence 46297, A
8	83	13.9	99	2 US-09-252-91A-17867 Sequence 17867, A
9	83	13.9	119	2 US-09-489-039A-10006 Sequence 10006, A
10	72	12.1	352	2 US-09-248-792B-17066 Sequence 17066, A
11	71.5	12.0	400	2 US-09-252-91A-22821 Sequence 22821, A
12	70	11.7	695	2 US-09-489-039A-14338 Sequence 14338, A
13	69	11.6	874	2 US-10-163-214-13 Sequence 14099, A
14	69	11.6	915	2 US-10-163-214-6 Sequence 6, Appli
15	68.5	11.5	409	2 US-09-232-993A-29381 Sequence 29381, A
16	68	11.4	43	1 US-09-433-241A-10 Sequence 10, Appli
17	68	11.4	735	2 US-09-115-704-2 Sequence 2, Appli
18	68	11.4	735	2 US-09-780-114-2 Sequence 2, Appli
19	68	11.4	904	2 US-10-163-214-2 Sequence 2, Appli
20	68	11.4	904	2 US-10-163-214-13 Sequence 2, Appli
21	67.5	11.3	1372	2 US-09-902-540-14099 Sequence 14099, A
22	67	11.2	299	2 US-09-720-318A-4 Sequence 4, Appli
23	67	11.2	696	1 US-09-081-55 Sequence 5, Appli
24	67	11.2	696	2 US-09-082-55 Sequence 5, Appli
25	67	11.2	696	4 PCT-US95-06934-5 Sequence 5, Appli
26	67	11.2	703	4 PCT-US95-06934-8 Sequence 8, Appli
27	11.2	718	4 PCT-US95-06934-6 Sequence 6, Appli	

#### ALIGNMENTS

RESULT 1	US-09-513-999C-7234	Sequence 739, APP
	; Sequence 7234, Application US/09513999C	Sequence 699, APP
	; Patent No. 6783961	Sequence 339, APP
	; GENERAL INFORMATION:	Sequence 339, APP
	; APPLICANT: Dumas Milne Edwards, J.B.	Sequence 339, APP
	; APPLICANT: Ducleit, A.	Sequence 339, APP
	; APPLICANT: Giordano, J.Y.	Sequence 339, APP
	; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.	Sequence 339, APP
	; Patent No. 6783961	Sequence 339, APP
	; IIE REFERENCE: 59, US2, REG	Sequence 339, APP
	; CURRENT APPLICATION NUMBER: US/09/513, 999C	Sequence 339, APP
	; PRIORITY APPLICATION NUMBER: US 60/122, 487	Sequence 339, APP
	; PRIORITY FILING DATE: 1999-02-26	Sequence 339, APP
	; NUMBER OF SEQ ID NOS: 36681	Sequence 339, APP
	; SOFTWARE: Patent, pm	Sequence 339, APP
	; SEQ ID NO: 7234	Sequence 339, APP
	; LENGTH: 90	Sequence 339, APP
	; TYPE: PRT	Sequence 339, APP
	; ORGANISM: Homo sapiens	Sequence 339, APP
	; FEATURE: NAME/KEY: UNSURE	Sequence 339, APP
	; LOCATION: 18	Sequence 339, APP
	; OTHER INFORMATION: Xaa=Asp or Glu	Sequence 339, APP
	; FEATURE: NAME/KEY: UNSURE	Sequence 339, APP
	; LOCATION: 46	Sequence 339, APP
	; OTHER INFORMATION: Xaa=Leu or Val	Sequence 339, APP
	; FEATURE: NAME/KEY: UNSURE	Sequence 339, APP
	; LOCATION: 78	Sequence 339, APP
	; OTHER INFORMATION: Xaa=Ala or Gly	Sequence 339, APP
	; US-09-513-999C-7234	Sequence 339, APP
	Query Match 67.5%; Score 403; DB 2; Length 90;	Sequence 739, APP
	Best Local Similarity 87.6%; Pred. No. 4.1e-40; Matches 78; Conservative 2; Mismatches 9; Index 0; Gaps 0;	Sequence 699, APP
	Oy	Sequence 339, APP
	1 MSGEPGQTSVAPPPEEVPG.....ASNGETLKKITNSRPPCVIL 115	Sequence 339, APP
	Db	1 MSGEPGQTSVAPPPEEVPG.....ASNGETLKKITNSRPPCVIL 115
	Db	1 MSGEPGQTSVAPPPEEVPG.....ASNGETLKKITNSRPPCVIL 115
	Oy	61 GTGAAFFIEINQOLVSKLNGGPFIEKDL 89
	Db	61 GTGAAFFIEINQOLVQXGNGFPYEVKDL 89

US-08-807-043-3	US-09-198-452A-739	Sequence 739, APP
	2 US-09-438-695A-699	Sequence 699, APP
	2 US-09-907-94A-339	Sequence 339, APP
	2 US-09-905-125A-339	Sequence 339, APP
	2 US-09-906-775A-339	Sequence 339, APP
	2 US-09-906-636-339	Sequence 339, APP
	2 US-09-904-62-339	Sequence 339, APP
	2 US-09-902-352A-339	Sequence 339, APP
	2 US-09-906-722A-339	Sequence 339, APP
	2 US-09-477-92A-101	Sequence 101, APP
	2 US-09-489-039A-7249	Sequence 7249, APP
	2 US-09-540-236-3416	Sequence 3416, APP

Sequence 3, Application US/08807043  
 Patent No. 5856311  
 GENERAL INFORMATION:  
 APPLICANT: Hillman, Jennifer L.  
 APPLICANT: Goli, Surya K.  
 TITLE OF INVENTION: NOVEL HUMAN SELENOPROTEIN  
 NUMBER OF SEQUENCES: 3  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Incyte Pharmaceuticals, Inc.  
 STREET: 3174 Porter Drive  
 CITY: Palo Alto  
 STATE: CA  
 COUNTRY: USA  
 ZIP: 94304  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Diskette  
 COMPUTER: IBM Compatible  
 OPERATING SYSTEM: DOS  
 SOFTWARE: FastSEQ for Windows Version 2.0  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/127,289  
 FILING DATE:  
 CLASSIFICATION:  
 PRIORITY APPLICATION DATA:  
 APPLICATION NUMBER: 08/807,043  
 FILING DATE:  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Billings, Lucy J.  
 REGISTRATION NUMBER: 36,749  
 REFERENCE/DOCKET NUMBER: PF-0202 US  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 415-845-4166  
 TELEX:  
 INFORMATION FOR SEQ ID NO: 3:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 88 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 IMMEDIATE SOURCE:  
 LIBRARY: GenBank  
 CLONE: 993035  
 US-08-807-043-3

Query Match 16.2%; Score 96.5; DB 1; Length 88;

Best Local Similarity 34.1%; Pred. No. 0.00087; Matches 28; Conservative 17; Mismatches 26; Indels 11; Gaps 5;

Qy 23 VRIVVEYCEPCGFEATYLEASAVKQYGP-IEI-ESRLGGTGAFFIEINGQLOVSKLE 79  
 Db 5 VRVY--YCGAXGKPKYLOKKEHEFPGLDICGEGTPQVTFPEVAGKLVHSKKR 62  
 Qy 80 NGCF-----PYEKDLIARRA 96  
 Db 63 GCGYVDTESKFRK-LVTAIKAA 83  
 US-09-127-289-3

RESULT 4

US-09-393-448-3  
 Sequence 3, Application US/09393448  
 Patent No. 6545129

GENERAL INFORMATION:  
 APPLICANT: Hillman, Jennifer L.  
 APPLICANT: Goli, Surya K.  
 TITLE OF INVENTION: NOVEL HUMAN SELENOPROTEIN  
 NUMBER OF SEQUENCES: 3  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Incyte Pharmaceuticals, Inc.  
 STREET: 3174 Porter Drive  
 CITY: Palo Alto  
 STATE: CA  
 COUNTRY: USA  
 ZIP: 94304  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Diskette  
 COMPUTER: IBM Compatible  
 OPERATING SYSTEM: DOS  
 SOFTWARE: FastSEQ for Windows Version 2.0  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/393,448  
 FILING DATE: 10-Sep-1999  
 CLASSIFICATION: <Unknown>  
 PRIORITY APPLICATION DATA:  
 APPLICATION NUMBER: 08/807,043  
 FILING DATE: <Unknown>



RESULT 7  
 US-09-393-448-1  
 ; Sequence 1, Application US/09393448  
 ; Patent No. 6595129  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Hillman, Jennifer L.  
 ; Goli, Surya K.  
 ; TITLE OF INVENTION: NOVEL HUMAN SELENOPROTEIN  
 ; NUMBER OF SEQUENCES: 3  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Incyte Pharmaceuticals, Inc.  
 ; STREET: 3174 Porter Drive  
 ; CITY: Palo Alto  
 ; STATE: CA  
 ; COUNTRY: USA  
 ; ZIP: 94304  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Diskette  
 ; COMPUTER: IBM Compatible  
 ; OPERATING SYSTEM: DOS  
 ; SOFTWARE: FastSeq for Windows Version 2.0  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/393,448  
 ; FILING DATE: 10-SEP-1999  
 ; CLASSIFICATION: <Unknown>  
 ; PRIORITY APPLICATION DATA:  
 ; APPLICATION NUMBER: 08/807,043  
 ; FILING DATE: <Unknown>  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Billing, Lucy J.  
 ; REGISTRATION NUMBER: 36,749  
 ; REFERENCE/DOCKET NUMBER: FF-0202 US  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 415-855-0555  
 ; TELEFAX: 415-845-4166  
 ; TELEX: <Unknown>  
 ; INFORMATION FOR SEQ ID NO: 1:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 93 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS: Single  
 ; TOPOLOGY: linear  
 ; IMMEDIATE SOURCE:  
 ; LIBRARY: BLADNOT03  
 ; CLONE: 1599862  
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
 ; US-09-393-448-1  
 ;  
 ; Query Match 15 2%; Score 96.5; DB 2; Length 93;  
 ; Best Local Similarity 32.8%; Pred. No. 0.00094;  
 ; Matches 21; Conservative 18; Mismatches 20; Indels 5; Gaps 3;  
 ;  
 ; QY 23 VRIVVYECSPCGFATEYLELLASAVKEQYQPG-IEI--ESRIGGTGAFETIEINGQLVFSKLE 79  
 ; 11 VRIVVYECSPCGFATEYLELLASAVKEQYQPG-IEI--ESRIGGTGAFETIEINGQLVFSKLE 68  
 ; QY 80 NGGF 83  
 ; 69 GDDY 72  
 ;  
 ; RESULT 8  
 ; US-09-270-767-46297  
 ; Sequence 46297, Application US/09270767  
 ; Patent No. 6703491  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Gary Breton et. al.  
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA  
 ; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS  
 ;  
 ; APPLICANT: Hamburger et. al.  
 ; TITLE OF INVENTION: Nucleic acids and proteins of *Drosophila melanogaster*  
 ; FILE REFERENCE: File Reference: 7326-094  
 ; CURRENT APPLICATION NUMBER: US/09/270,767  
 ; CURRENT FILING DATE: 1999-03-17  
 ; NUMBER OF SEQ ID NOS: 62217  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO: 46297  
 ; LENGTH: 216  
 ; TYPE: PRT  
 ; ORGANISM: *Drosophila melanogaster*  
 ; US-09-270-767-46297  
 ;  
 ; Query Match 14.6%; Score 87; DB 2; Length 216;  
 ; Best Local Similarity 18.8%; Pred. No. 0.042; Matches 25; Conservative 15; Mismatches 25; Indels 68; Gaps 1;  
 ;  
 ; QY 29 YCERPGCFETMYLELASAVKQYQPGTEI- 55  
 ; 66 YCYSGYRAFEDTVGLGSKYQPOIQVNGCNYDPGPNLYLSKMFALAKLIIIVSVA 125  
 ; QY 56 -----ESRIGGTGAFETIEINGQLVFSKLE 80  
 ; Db 126 SPFTFLGLNTPSWNSHMQANKIYACMMIFLGLNMLEAQOLISSGAFETITNDVPSKLOT 185  
 ; Db 186 GRFPSPPEVLFQII 198  
 ;  
 ; RESULT 9  
 ; US-09-252-991A-17867  
 ; Sequence 17867, Application US/09252991A  
 ; Patent No. 6551795  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Marc J. Rubenfield et al.  
 ; TITLE OF INVENTION: NUCICIC ACID AND AMINO ACID SEQUENCES RELATING TO *PSEUDOMONAS*  
 ; FILE REFERENCE: 10196-136  
 ; CURRENT APPLICATION NUMBER: US/09/252,991A  
 ; CURRENT FILING DATE: 1999-02-18  
 ; PRIOR APPLICATION NUMBER: US 60/074,788  
 ; PRIOR FILING DATE: 1998-02-18  
 ; PRIOR APPLICATION NUMBER: US 60/094,190  
 ; PRIOR FILING DATE: 1998-07-27  
 ; NUMBER OF SEQ ID NOS: 33142  
 ; SEQ ID NO: 17867  
 ; LENGTH: 99  
 ; TYPE: PRT  
 ; ORGANISM: *Pseudomonas aeruginosa*  
 ; US-09-252-991A-17867  
 ;  
 ; Query Match 13.9%; Score 83; DB 2; Length 99;  
 ; Best Local Similarity 31.2%; Pred. No. 0.041; Matches 25; Conservative 13; Mismatches 36; Indels 6; Gaps 3;  
 ;  
 ; QY 19 PGSGVRIVVYECSPCGF--EATLIL--ELMSAVKEQYQPGTEIERSRIGGTGAFETIEINGQLV 74  
 ; Db 5 PTKARPEIVITVYTCQWMLRAAQLQELSTFADDLGKVCLEPCGTG-GVFRITCDGVQV 62  
 ; QY 75 FSKLENGGPYKEOLIEAR 94  
 ; Db 63 WERKADGGPPEAKKQRVR 82  
 ;  
 ; RESULT 10  
 ; US-09-439-039A-10006  
 ; Sequence 10006, Application US/09489039A  
 ; Patent No. 6610836  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Gary Breton et. al.  
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA  
 ; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS



; LENGTH: 874  
; TYPE: PRT  
; ORGANISM: *Vigna unguiculata*  
; US-10-163-214-13

Query Match 11.6%; Score 69; DB 2; Length 874;  
 Best Local Similarity 26.8%; Pred. No. 43;  
 Matches 22; Conservative 14; Mismatches 34; Indels 12; Gaps 3;  
 QY 21 SGVRIVVEYCEPCFEATYLELASVAKEDQYQPGTIESRLGGTGAFFIEINGQLVFSKLEN 80  
 Db 754 SDMFIPIFSIPEG-----ITQMSMRYGAQPIARKTGGLNSVFDVDDTISQFRN 806  
 QY 81 GGFPY---EKDLIEAIRRASN 98  
 Db 807 -GPTFLNADEKGINDALVRAIN 827

RESULT 15

US-10-163-214-6  
 Sequence 6, Application US/10163214  
 ; Patent No. 6849781  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Allen, Stephen M.  
 ; APPLICANT: Brodie, Karen E.  
 ; APPLICANT: Butler, Karlene H.  
 ; APPLICANT: Thorpe, Catherine J.  
 ; TITLE OF INVENTION: Starch Synthase Isoform V  
 ; FILE REFERENCE: BB1120 US NA  
 ; CURRENT APPLICATION NUMBER: US/10/163,214  
 ; CURRENT FILING DATE: 2005-06-05  
 ; PRIOR APPLICATION NUMBER: 60/297,099  
 ; PRIOR FILING DATE: 2001-06-08  
 ; NUMBER OF SEQ ID NOS: 13  
 ; SOFTWARE: Microsoft Office 97  
 ; SEQ ID NO 6  
 ; LENGTH: 915  
 ; TYPE: PRT  
 ; ORGANISM: *Oryza sativa*  
 ; US-10-163-214-6

Query Match 11.6%; Score 69; DB 2; Length 915;  
 Best Local Similarity 28.4%; Pred. No. 46;  
 Matches 23; Conservative 11; Mismatches 37; Indels 10; Gaps 2;  
 QY 21 SGVRIVVEYCEPCFEATYLELASVAKEDQYQPGTIESRLGGTGAFFIEINGQLVFSKLEN 80  
 Db 797 SDMFIPIFSIPEG-----ITQMSMRYGAQPIARKTGGLNSVFDVDDTISQFRN 849  
 QY 81 GGFPY---EKDLIEAIRRASN 98  
 Db 850 GFTFVHPDKEKLSGAMERAFN 870

Search completed: March 27, 2006, 12:10:05  
 Job time : 30.951 secs

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Biocceleration Ltd.

## OM protein - protein search, using sw model

Run on: March 27, 2006, 12:32:34 ; Search time 12.7395 Seconds  
(without alignments)  
266.247 Million cell updates/sec

Title: US-09-824-787B-2

Perfect score: 597

Sequence: 1 MSGEPGQTSVAPPPEVEPG.....ASNGETLEKINTNSRPPCVIL 115

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 174695 seqs, 29454374 residues

Total number of hits satisfying chosen parameters: 174695

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Database : Listing first 45 summaries

1: Published Applications AA New:  
1: /SIDSS5ptodata/2/pubpaa/US05\_NEW\_PUB.pep:  
2: /SIDSS5ptodata/2/pubpaa/US07\_NEW\_PUB.pep:  
3: /SIDSS5ptodata/2/pubpaa/PCT\_NEW\_PUB.pep:  
4: /SIDSS5ptodata/2/pubpaa/US05\_NEW\_PUB.pep:  
5: /SIDSS5ptodata/2/pubpaa/US10\_NEW\_PUB.pep:  
6: /SIDSS5ptodata/2/pubpaa/US11\_NEW\_PUB.pep:  
7: /SIDSS5ptodata/2/pubpaa/US60\_NEW\_PUB.pep:  
8: /SIDSS5ptodata/2/pubpaa/US60\_NEW\_PUB.pep:  
;

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	597	100.0	115	7 US-11-155-288-11
2	76.5	12.8	170	7 US-11-096-568A-1010
3	76.5	12.8	186	7 US-11-096-568A-10109
4	76.5	12.8	198	7 US-11-096-568A-10109
5	76	12.7	283	7 US-11-096-568A-31233
6	76	12.7	390	7 US-11-096-568A-31232
7	76	12.7	395	7 US-11-096-568A-31231
8	69	11.6	253	7 US-11-096-568A-25229
9	69	11.6	67	7 US-11-096-568A-25228
10	69	11.6	415	7 US-11-096-568A-25227
11	68.5	11.5	186	7 US-11-096-568A-22856
12	68.5	11.5	209	7 US-11-096-568A-22855
13	68.5	11.5	234	7 US-11-096-568A-22854
14	68.5	11.5	329	6 US-10-510-386-36
15	68.5	11.5	330	6 US-10-510-386-196
16	67	11.2	228	7 US-11-082-389-24
17	66.5	11.1	697	7 US-11-096-568A-28319
18	66.5	11.1	759	6 US-11-047-657-3280
19	66.5	11.1	876	7 US-11-096-568A-28318
20	66.5	11.1	884	7 US-11-096-568A-28317
21	66	11.1	244	5 US-09-978-3608-510
22	66	11.1	259	7 US-11-096-568A-21634
23	66	11.1	596	7 US-11-152-903-2
24	66	11.1	596	7 US-11-152-903-4
25	66	11.1	596	7 US-11-152-903-6

## ALIGNMENTS

RESULT 1  
US-11-155-288-11  
; Sequence 11, Application US/11155288  
; Publication No. US20060008468A1  
; GENERAL INFORMATION:  
; APPLICANT: Chiang, Chih-Sheng  
; APPLICANT: Simard, John J.L.  
; TITLE OF INVENTION: COMBINATIONS OF TUMOR ASSOCIATED  
; FILE REFERENCE: MANK.050A  
; CURRENT APPLICATION NUMBER: US/11/155, 288  
; CURRENT FILING DATE: 2005-06-17  
; PRIOR APPLICATION NUMBER: 60/580, 969  
; PRIOR FILING DATE: 2004-06-17  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; NUMBER OF SEQ ID NOS: 40  
; SEQ ID NO 11  
; LENGTH: 115  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; US-11-155-288-11  
Query Match Best Local Similarity 100.0%; Pred. No. 4. 8e-58; Score 597; DB 7; Length 115; Matches 115; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 MSGEPGQTSVAPPPEVEPG.....ASNGETLEKINTNSRPPCVIL 115  
Db 61 GTGAFETEINGOLVFSKLENGGFPYKEDLIEARRASNGETLEKINTNSRPPCVIL 115  
; RESULT 2  
US-11-096-568A-1010  
; Sequence 10110, Application US/11096568A  
; Publication No. US20060048240A1  
; GENERAL INFORMATION:  
; APPLICANT: Alexandrov, Nickolai et al.  
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides  
; TITLE OF INVENTION: Therry  
; FILE REFERENCE: 2750-1592PUS2  
; CURRENT APPLICATION NUMBER: US/11/096, 568A  
; CURRENT FILING DATE: 2005-04-01  
; NUMBER OF SEQ ID NOS: 34471  
; SEQ ID NO 10110

LENGTH: 170  
 TYPE: PRT  
 ORGANISM: *Triticum aestivum*  
 FEATURE: misc feature  
 NAME/KEY: misc feature  
 LOCATION: (1)..(170)  
 OTHER INFORMATION: Ceres Seq. ID no. 13583721  
 US-11-096-568A-101010

Query Match 12.8%; Score 76.5; DB 7; Length 170;  
 Best Local Similarity 31.4%; Pred. No. 0.47; Mismatches 11; Indels 3; Gaps 2;  
 Matches 22; Conservative 11; Mismatches 34; Indels 3; Gaps 2;

Qy 21 SGVRIVIVEYCEPC-GFREATYLELASAVQYQGIE--IESRLGQTGAFETEINGOLVFSK 77  
 Db 83 SGKRIVIVEACTQCOQFKRRAKVKEDLESAVPGVSVTINPEKPRRGCLEIREEGDVFIS 142

Qy 78 LENGGPPEYK 87  
 Db 143 LQNMPRPFKK 152

RESULT 3  
 US-11-096-568A-10109  
 Sequence 10109, Application US/11096568A  
 Publication No. US20060048240A1  
 GENERAL INFORMATION:  
 APPLICANT: Alexandrov, Nickolai et al.  
 TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides  
 TITLE OF INVENTION: Therry  
 FILE REFERENCE: 2750-1592PUS2  
 CURRENT APPLICATION NUMBER: US/11/096,568A  
 CURRENT FILING DATE: 2005-04-01  
 NUMBER OF SEQ ID NOS: 34471  
 SEQ ID NO 10109  
 LENGTH: 186  
 TYPE: PRT  
 ORGANISM: *Triticum aestivum*  
 FEATURE:  
 NAME/KEY: misc feature  
 LOCATION: (1)..(186)  
 OTHER INFORMATION: Ceres Seq. ID no. 13583720  
 US-11-096-568A-10109

Query Match 12.8%; Score 76.5; DB 7; Length 186;  
 Best Local Similarity 31.4%; Pred. No. 0.53; Mismatches 11; Indels 3; Gaps 2;  
 Matches 22; Conservative 11; Mismatches 34; Indels 3; Gaps 2;

Qy 21 SGVRIVIVEYCEPC-GFREATYLELASAVQYQGIE--IESRLGQTGAFETEINGOLVFSK 77  
 Db 99 SGKRIVIVEACTQCOQFKRRAKVKEDLESAVPGVSVTINPEKPRRGCLEIREEGDVFIS 158

Qy 78 LENGGPPEYK 87  
 Db 159 LQNMPRPFKK 168

RESULT 4  
 US-11-096-568A-10108  
 Sequence 10108, Application US/11096568A  
 Publication No. US20060048240A1  
 GENERAL INFORMATION:  
 APPLICANT: Alexandrov, Nickolai et al.  
 TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides  
 TITLE OF INVENTION: Therry  
 FILE REFERENCE: 2750-1592PUS2  
 CURRENT APPLICATION NUMBER: US/11/096,568A  
 CURRENT FILING DATE: 2005-04-01  
 NUMBER OF SEQ ID NOS: 34471  
 SEQ ID NO 10108  
 LENGTH: 198  
 TYPE: PRT  
 ORGANISM: *Triticum aestivum*

RESULT 5  
 US-11-096-568A-31233  
 Sequence 31233, Application US/11096568A  
 Publication No. US20060048240A1  
 GENERAL INFORMATION:  
 APPLICANT: Alexandrov, Nickolai et al.  
 TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides  
 FILE REFERENCE: 2750-1592PUS2  
 CURRENT APPLICATION NUMBER: US/11/096,568A  
 CURRENT FILING DATE: 2005-04-01  
 NUMBER OF SEQ ID NOS: 34471  
 SEQ ID NO 31233  
 LENGTH: 283  
 TYPE: PRT  
 ORGANISM: *Arabidopsis thaliana*  
 FEATURE:  
 NAME/KEY: misc feature  
 LOCATION: (1)..(283)  
 OTHER INFORMATION: Ceres Seq. ID no. 13579073  
 US-11-096-568A-31233

Query Match 12.7%; Score 76; DB 7; Length 283;  
 Best Local Similarity 23.3%; Pred. No. 1; Mismatches 31; Conservative 16; Mismatches 36; Indels 50; Gaps 5;  
 Matches 31; Conservative 16; Mismatches 36; Indels 50; Gaps 5;

Qy 13 PPEEVEP-----GSVRIVIVEYCEPC-GFREATYLELASAVQYQGIE--IESRL-- 59  
 Db 96 PPNKARPYDVLGAALFGDPAAVITGADRECEAPFMEHYAVQQLPSPVNDGRLE 155

Qy 60 -----GTTGAFFIEIN-----GOLVFSKLENGSF 83  
 Db 156 EGINFKLGLDPQKIEENIEFCIKLMLGAGDESMEFNMFWAVHPPGAILNRLT-KL 214

Qy 84 PYKDQLEAIRRA 96  
 Db 215 KLEKEKJESRRRA 227

RESULT 6  
 US-11-096-568A-31232  
 Sequence 31232, Application US/11096568A  
 Publication No. US20060048240A1  
 GENERAL INFORMATION:  
 APPLICANT: Alexandrov, Nickolai et al.  
 TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides  
 TITLE OF INVENTION: Therry  
 FILE REFERENCE: 2750-1592PUS2  
 CURRENT APPLICATION NUMBER: US/11/096,568A  
 CURRENT FILING DATE: 2005-04-01  
 NUMBER OF SEQ ID NOS: 34471  
 SEQ ID NO 31232  
 LENGTH: 390  
 TYPE: PRT

ORGANISM: *Arabidopsis thaliana*  
 FEATURE: misc\_feature  
 NAME/KEY: misc\_feature  
 LOCATION: (1)..(310)  
 OTHER INFORMATION: Ceres Seq. ID no. 13579072  
 US-11-096-568A-31232

Query Match 12.7%; Score 76; DB 7; Length 390;  
 Best Local Similarity 23.3%; Pred. No. 1.6; Mismatches 16; Indels 36; Gaps 5;  
 Matches 31; Conservative 16; MisMatches 36; Indels 50; Gaps 5;

QY 13 PPRVEP-----GSGVRIVVEYCEPCGFATYLELASAVKEQYQGIE-IESR-- 59  
 Db 203 PPKKARPDLVGAALFGDGAAGAVIGADPRECEAPFMLEHYAVQOFLPGTQNYIDGRITE 262  
 QY 60 -----GCTGAFBIEIN-----GOLVFSKLENGF 83  
 Db 263 EGINFKULGRDLPKQKIEENIEFCKKLMGSKADESMEFNDMFWAVHPGGPAINRLET-KL 321  
 QY 84 PYEKLDEBAIRRA 96  
 Db 322 KLEKEKLEKSSRRA 334

RESULT 7  
 US-11-096-568A-31231  
 Sequence 31231; Application US/11096568A  
 Publication No. US20060048240A1  
 GENERAL INFORMATION:  
 APPLICANT: Alexandrov, Nickolai et al.  
 TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides  
 TITLE OF INVENTION: Therry  
 FILE REFERENCE: 2750-1592PUS2  
 CURRENT APPLICATION NUMBER: US/11/096,568A  
 CURRENT FILING DATE: 2005-04-01  
 NUMBER OF SEQ ID NOS: 34471  
 SEQ ID NO 31231  
 LENGTH: 395

TYPE: PRT  
 ORGANISM: *Arabidopsis thaliana*  
 FEATURE:  
 NAME/KEY: misc\_feature  
 LOCATION: (1)..(395)  
 OTHER INFORMATION: Ceres Seq. ID no. 13579071  
 US-11-096-568A-31231

Query Match Best Local Similarity 12.3%; Score 76; DB 7; Length 395;  
 Matches 31; Conservative 16; Mismatches 36; Indels 50; Gaps 5;

QY 13 PPRVEP-----GSGVRIVVEYCEPCGFATYLELASAVKEQYQGIE-IESR-- 59  
 Db 208 PPKNKARPDLVGAALFGDGAAGAVIGADPRECEAPFMLEHYAVQOFLPGTQNYIDGRITE 267  
 QY 60 -----GCTGAFBIEIN-----GOLVFSKLENGF 83  
 Db 268 EGINFKULGRDLPKQKIEENIEFCKKLMGSKADESMEFNDMFWAVHPGGPAINRLET-KL 326  
 QY 84 PYEKLDEBAIRRA 96  
 Db 327 KLEKEKLEKSSRRA 339

RESULT 8  
 US-11-096-568A-25229  
 Sequence 25229; Application US/11096568A  
 Publication No. US20060048240A1  
 GENERAL INFORMATION:  
 APPLICANT: Alexandrov, Nickolai et al.  
 TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides  
 TITLE OF INVENTION: Therry  
 FILE REFERENCE: 2750-1592PUS2  
 CURRENT APPLICATION NUMBER: US/11/096,568A

RESULT 9  
 US-11-096-568A-25228  
 Sequence 25228; Application US/11096568A  
 Publication No. US20060048240A1  
 GENERAL INFORMATION:  
 APPLICANT: Alexandrov, Nickolai et al.  
 TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides  
 TITLE OF INVENTION: Therry  
 FILE REFERENCE: 2750-1592PUS2  
 CURRENT APPLICATION NUMBER: US/11/096,568A  
 CURRENT FILING DATE: 2005-04-01  
 NUMBER OF SEQ ID NOS: 34471  
 SEQ ID NO 25228  
 LENGTH: 367

TYPE: PRT  
 ORGANISM: *Zea mays* subsp. *mays*  
 FEATURE:  
 NAME/KEY: misc\_feature  
 LOCATION: (1)..(367)  
 OTHER INFORMATION: Ceres Seq. ID no. 12566622  
 US-11-096-568A-25228

Query Match Best Local Similarity 11.6%; Score 69; DB 7; Length 367;  
 Matches 29; Conservative 17; Mismatches 45; Indels 30; Gaps 5;

QY 12 PPRVEVERGSGRIV-----VCEC-----PGCFATYLELASAVKEQYQGIEISR 58  
 Db 137 PPKDPLPINTALAVASWGVYVWLTSDRDLPGGSSHFAOTVKAELKKGILVE-- 194  
 QY 59 LGCTGAFBIEINQOL-VFSKLENGFPEKDLIEAIR-----RASNGETLEKINTS 108  
 Db 195 -----CLTSDFRGLDEAISLANSGLDVYAHMIEITVRSQVRDPRAGYDQSIAVLKHA 249  
 QY 109 R 109  
 Db 250 K 250

RESULT 10  
 US-11-096-568A-25227  
 Sequence 25227; Application US/11096568A  
 Publication No. US20060048240A1  
 GENERAL INFORMATION:

APPLICANT: Alexandrov, Nickolai et al.  
 TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides  
 FILE REFERENCE: 2750-1592PUS2  
 CURRENT APPLICATION NUMBER: US/11/096,568A  
 CURRENT FILING DATE: 2005-04-01  
 NUMBER OF SEQ ID NOS: 34471  
 SEQ ID NO: 25227  
 LENGTH: 415  
 TYPE: PRT  
 ORGANISM: Zea mays subsp. mays  
 FEATURE:  
 NAME/KEY: misc\_feature  
 LOCATION: (1)..(415)  
 OTHER INFORMATION: Ceres Seq. ID no. 12566621  
 ;US-11-096-568A-25227

Query Match 11.6%; Score 69; DB 7; Length 415;  
 Best Local Similarity 24.0%; Pred. No. 10;  
 Matches 29; Conservative 17; Mismatches 45; Indels 30; Gaps 5;  
 QY 12 PPPEEVPGSQRIV---VERCE-----PGCFEATYLELASAVKQYQPGIEISR 58  
 Db 185 PPPDPPEPLNVALAVASWGVNQVVLTSVDRDDLPDGSSHRAQTWALKLKGIVLIVE-- 242  
 QY 59 LGGTGAFEIEINGOL-VFSKLENGGPFYKEDLIEAIR-----RASNGETLEKINTS 108  
 Db 243 ----CLTSDPFRGLEAISSLANSGLDGVVAYHNIETVRSQGRVVRDPRAGYDQSLAVLKA 297  
 QY 109 R 109  
 Db 298 K 298

RESULT 11  
 US-11-096-568A-22856  
 ; Sequence 22856, Application US/11096568A  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Alexandrov, Nikolai et al.  
 ; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides  
 ; TITLE OF INVENTION: Therby  
 ; FILE REFERENCE: 2750-1592US2  
 ; CURRENT APPLICATION NUMBER: US/11/096,568A  
 ; CURRENT FILING DATE: 2005-04-01  
 ; NUMBER OF SEQ ID NOS: 34471  
 ; SEQ ID NO: 22856  
 ; LENGTH: 186  
 ; TYPE: PRT  
 ; ORGANISM: Zea mays subsp. mays  
 ; FEATURE:  
 ; NAME/KEY: misc\_feature  
 ; LOCATION: (1)..(186)  
 ; OTHER INFORMATION: Ceres Seq. ID no. 12410397  
 ;US-11-096-568A-22856

Query Match 11.5%; Score 68.5; DB 7; Length 186;  
 Best Local Similarity 28.6%; Pred. No. 3.9;  
 Matches 28; Conservative 10; Mismatches 37; Indels 23; Gaps 4;  
 QY 3 GPGQTSVAPPPEPE---PGSGRIVIVYCEPGCFEATYLELASAVKQYQPGIEISR 58  
 Db 77 GKKEAEAVAPAAEEDBGDVDAAVGKRVIVAEACTC---RQFKIRAQKVE---DLSY 128  
 QY 59 LGGT-----GAFEIEINGOLVFSKLENGGPFY 85  
 Db 129 VSGVSVIINPQKPRRGCLEIREEGEVFISLNMPPF 166

RESULT 12  
 US-11-096-568A-22855  
 ; Sequence 22855, Application US/11096568A  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Andersen, Jens Toine

RESULT 13  
 US-11-096-568A-22854  
 ; Sequence 22854, Application US/11096568A  
 ; Publication No. US20060048240A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Alexandrov, Nickolai et al.  
 ; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides  
 ; TITLE OF INVENTION: Therby  
 ; FILE REFERENCE: 2750-1592PUS2  
 ; CURRENT APPLICATION NUMBER: US/11/096,568A  
 ; CURRENT FILING DATE: 2005-04-01  
 ; NUMBER OF SEQ ID NOS: 34471  
 ; SEQ ID NO: 22854  
 ; LENGTH: 234  
 ; TYPE: PRT  
 ; ORGANISM: Zea mays subsp. mays  
 ; FEATURE:  
 ; NAME/KEY: misc\_feature  
 ; LOCATION: (1)..(234)  
 ; OTHER INFORMATION: Ceres Seq. ID no. 12410395  
 ;US-11-096-568A-22854

Query Match 11.5%; Score 68.5; DB 7; Length 234;  
 Best Local Similarity 28.6%; Pred. No. 5.3;  
 Matches 28; Conservative 10; Mismatches 37; Indels 23; Gaps 4;  
 QY 3 GPGQTSVAPPPEPE---PGSGRIVIVYCEPGCFEATYLELASAVKQYQPGIEISR 58  
 Db 125 GKKEAEAVAPAAEEDBGDVDAAVGKRVIVAEACTC---RQFKIRAQKVE---DLSY 176  
 QY 59 LGGT-----GAFEIEINGOLVFSKLENGGPFY 85  
 Db 177 VSGVSVIINPQKPRRGCLEIREEGEVFISLNMPPF 214

RESULT 14  
 US-10-510-386-36  
 ; Sequence 36, Application US/10510386  
 ; Publication No. US20050244922A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Clausen, Ib Groth



Copyright (c) 1993 - 2006 Biocceleration Ltd.  
GenCore version 5.1.7

Om protein - protein search, using sw model  
Run on: March 27, 2006, 11:54:20 ; Search time 119.82 Seconds  
421.703 Million cell updates/sec

Title: US-09-824-787B-2  
Perfect score: 597

Sequence: 1 MSGEPGQTSVAPPPEVEPG.....ASNGETLEKITNSRPPCVIL 115,  
Scoring table: BLOSUM62

Searched: Gapext 10.0 , Gapext 0.5  
Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database : A\_Geneseq 21;\*  
1: geneseqD1980s;\*  
2: geneseqD1990s;\*  
3: geneseqD2000s;\*  
4: geneseqD2001s;\*  
5: geneseqD2002s;\*  
6: geneseqD2003as;\*  
7: geneseqD2003bs;\*  
8: geneseqD2004as;\*  
9: geneseqD2005s;\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query	Match %	Length	DB	ID	Description
1	597	100.0	115	4	AAG78997	Aag78997	Human C35
2	597	100.0	115	4	AAG77870	Aag77870	Human C35
3	597	100.0	115	5	ABP58560	Abp58560	Human sit
4	597	100.0	115	5	ABP43843	Abp43843	RIKEN 181
5	597	100.0	115	8	ADH13244	Adh13244	Human mall
6	597	100.0	115	8	ADK48916	Adk48916	Human bre
7	597	100.0	115	9	ADV60511	Adv60511	Human bre
8	597	100.0	115	9	AD813703	Ad813703	Human C35
9	597	100.0	115	9	AE015131	Aea015131	Human pol
10	597	100.0	115	9	AE02447	Aea02447	Human C35
11	597	100.0	115	5	ABP43055	Abp43055	Human ova
12	597	100.0	131	3	ABP43521	Abp43521	Human can
13	597	100.0	131	8	AD49069	Ad49069	Human bre
14	597	100.0	149	9	AE62239	Aea62239	Recombi
15	597	100.0	206	6	ABP47619	Abp47619	Breast ca
16	466.5	78.1	207	8	ADP84574	Adp84574	Human bre
17	466.5	78.1	208	8	ADP4574	Adp4574	Human bre
18	403	67.5	90	3	AA03153	Aaa03153	Human sec
19	403	67.5	90	8	ADX49067	Adx49067	Human bre
20	292	48.9	75	8	ADK50220	Adk50220	Human car
21	291	48.7	65.	8	ADK50219	Adk50219	Human cat
22	290	48.6	65	8	ADK50217	Adk50217	Human car
23	290	48.6	75	8	ADK50218	Adk50218	Human car
24	289	48.4	64	8	ADK50201	Adk50201	Human car

Result No.	Score	Query	Match %	Length	DB	ID	Description
1	597	100.0	115	4	AAG78997	Aag78997	Human C35
2	597	100.0	115	4	AAG77870	Aag77870	Human C35
3	597	100.0	115	5	ABP58560	Abp58560	Human sit
4	597	100.0	115	5	ABP43843	Abp43843	RIKEN 181
5	597	100.0	115	8	ADH13244	Adh13244	Human mall
6	597	100.0	115	8	ADK48916	Adk48916	Human bre
7	597	100.0	115	9	ADV60511	Adv60511	Human bre
8	597	100.0	115	9	AD813703	Ad813703	Human C35
9	597	100.0	115	9	AE015131	Aea015131	Human pol
10	597	100.0	115	9	AE02447	Aea02447	Human C35
11	597	100.0	115	5	ABP43055	Abp43055	Human ova
12	597	100.0	131	3	ABP43521	Abp43521	Human can
13	597	100.0	131	8	AD49069	Ad49069	Human bre
14	597	100.0	149	9	AE62239	Aea62239	Recombi
15	597	100.0	206	6	ABP47619	Abp47619	Breast ca
16	466.5	78.1	207	8	ADP84574	Adp84574	Human bre
17	466.5	78.1	208	8	ADP4574	Adp4574	Human bre
18	403	67.5	90	3	AA03153	Aaa03153	Human sec
19	403	67.5	90	8	ADX49067	Adx49067	Human bre
20	292	48.9	75	8	ADK50220	Adk50220	Human car
21	291	48.7	65.	8	ADK50219	Adk50219	Human cat
22	290	48.6	65	8	ADK50217	Adk50217	Human car
23	290	48.6	75	8	ADK50218	Adk50218	Human car
24	289	48.4	64	8	ADK50201	Adk50201	Human car

#### ALIGNMENTS

25	269	48.4	73	8	ADK50214	Adk50214	Human car
26	289	48.4	73	8	ADK50202	Adk50202	Human car
27	288	48.2	64	8	ADK50187	Adk50187	Human car
28	288	48.2	73	8	ADK50133	Adk50133	Human car
29	288	48.2	73	8	ADK50188	Adk50188	Human car
30	288	48.2	75	8	ADK50196	Adk50196	Human car
31	287	48.1	64	8	ADK50209	Adk50209	Human car
32	287	48.1	73	8	ADK50215	Adk50215	Human car
33	287	48.1	73	8	ADK50210	Adk50210	Human car
34	287	48.1	73	8	ADK50216	Adk50216	Human car
35	286	47.9	64	8	ADK50227	Adk50227	Human car
36	286	47.9	64	8	ADK50221	Adk50221	Human car
37	286	47.9	66	8	ADK50225	Adk50225	Human car
38	286	47.9	73	8	ADK50222	Adk50222	Human car
39	286	47.9	73	8	ADK50182	Adk50182	Human car
40	286	47.9	73	8	ADK50228	Adk50228	Human car
41	286	47.9	77	8	ADK50226	Adk50226	Human car
42	285	47.8	85	8	ADK50206	Adk50206	Human car
43	285	47.7	64	8	ADK50189	Adk50189	Human car
44	285	47.7	65	8	ADK50230	Adk50230	Human car

Query Match % Score Pred. No. 1.3e-60;  
Best Local Similarity 100.0%; Length 115;  
The present sequence is human C35. C35 is a novel tumour antigen that is overexpressed in human breast and bladder carcinoma. C35 is thought to be a promising candidate for tumour immunotherapy, in immunogenic compositions and vaccines, to induce antibody and cell-mediated immunity against target cells such as tumour cells that express C35 genes. The C35 gene aligns on human chromosome 17q12. Sequence 115 AA;



QY	61	GTGAFIEIENGQJLVPFKLENGGPPYKDLIEAIRRASNGETLEKINTNSRPCVIL	115	Db	61	GTGAFIEIENGQJLVPFKLENGGPPYKDLIEAIRRASNGETLEKINTNSRPCVIL	115	
Db	61	GTGAFIEIENGQJLVPFKLENGGPPYKDLIEAIRRASNGETLEKINTNSRPCVIL	115	RESULT 4				
ID	ABP43843	standard; protein; 115 AA.		RESULT 5				
XX				ID	ADH13244	standard; protein; 115 AA.		
AC	ABP43843;			ID	ADH13244			
XX				AC	ADH13244;			
DT	26-FEB-2003	(first entry)		XX				
XX				DT	11-MAR-2004	(first entry)		
DE	RIKEN 1810046J19 protein.			XX				
XX				DE	Human malignant neoplasia-related protein SeqID93.			
KW	Neuroprotective; immunomodulator; cancer; chromosome 4q13-q21; cytostatic; anti-inflammatory; gene therapy; Huntington's disease; wound; burn; ulcer; Alzheimer's disease; amyotrophic lateral sclerosis; autoimmune disorder; inflammation; vulnerability.			XX				
KW				KW	malignant neoplasia; cytostatic; breast cancer; ovarian cancer; gastric cancer; colon cancer; oesophageal cancer; mesenchymal cancer; bladder cancer; non-small cell lung cancer; human.			
KW				XX				
OS	Homo sapiens.			OS	Homo sapiens.			
XX				XX				
PN	WC200231111-A2.			PN	EP1365034-A2.			
PD	18-APR-2002.			XX				
PF	11-OCT-2001; 2001WO-US027760.			PD	26-NOV-2003.			
XX				XX				
PR	12-OCT-2000; 2000US-00687527.			PF	09-MAY-2003; 2003EP-00010447.			
XX				XX				
PA	(HYSE-) HYSEQ INC.			PR	21-MAY-2002; 2002EP-00010291.			
PI	Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;			XX	13-FEB-2003; 2003EP-00003112.			
PI	Xue AJ, Yang Y, Wehrman T, Drmanac RT;			XX				
XX				PA	(FARB ) BAYER AG.			
DR	WPI; 2002-426278/45. N-PSDB; A8061087.			PA				
XX				PI	Wirtz R, Munnes M, Kallabis H;			
PA				XX				
XX				DR				
PT	New polypeptides and their encoded proteins, useful as nutritional sources or supplements, or in gene therapy, particularly for treating wounds, Alzheimer's disease, amyotrophic lateral sclerosis, cancer or PT inflammation.			XX				
XX				PT	Predicting, diagnosing or prognosing malignant neoplasia by detecting at least two markers, where the markers are genes from one or more chromosomal regions altered in malignant neoplasia.			
PS	Claim 20; SEQ ID # 746; 357pp + Sequence Listing; English.			PT				
XX				PS	Claim 11; SEQ ID NO 93; 267pp; English.			
CC	The invention relates to 446 newly isolated polynucleotide sequences. The activity of polynucleotides of the invention may be described as:			XX				
CC	vulnary, neuroprotective, immunomodulator, cytostatic and anti-inflammatory. Compositions comprising nucleic acids of the invention are useful for treating a mammal subiect, or as nutritional sources or supplements. These are useful in gene therapy, particularly for treating wounds, burns or ulcers, Alzheimer's disease, amyotrophic lateral sclerosis, autoimmune disorders, cancer or inflammation. The nucleic acids and polypeptides are also useful in diagnostic and research methods. The sequences given in records ABP43544-ABP43899 represent polypeptides encoded by polynucleotides of the invention. NOTE: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WPO at <a href="ftp://wipo.int/pdb/published_pct_sequences">ftp://wipo.int/pdb/published_pct_sequences</a>			CC				
CC	XX			CC	This invention relates to a novel method for the prediction, diagnosis, or prognosis of malignant neoplasia by the detection of at least two markers. The invention may also be useful for the development of cytostatic compounds through the regulation of the expression of a gene or activity of a protein associated with malignant neoplasia. The method is useful for prediction, diagnosis or prognosis of malignant neoplasia such as breast cancer, ovarian cancer, gastric cancer, colon cancer, oesophageal cancer, mesenchymal cancer, bladder cancer or non-small cell lung cancer. The polynucleotides and polypeptides defined in the specification, antisense polynucleotides targeting the polynucleotides, antibodies targeting either one of the polynucleotides or polypeptides, and compounds identified by the screening methods are useful for preventing or treating malignant neoplasia. The disease treated is preferably breast cancer. The present sequence is that of a human malignant neoplasia-related protein which may be used in the method of the invention.			
CC	XX			CC	Sequence 115 AA;			
SQ	Sequence 115 AA;			QY	Query Match 100.0%; Score 597; DB 8; Length 115;			
				Db	Best Local Similarity 100.0%; Pred. No. 1.3e-60; Matches 115; Conservative 0; Mismatches 0; Indels 0; Gaps 0.			
Query Match	100.0%; Score 597; DB 5; Length 115;			QY	1 MSGEQQTSAVAPPPEVEPGSGRIVVEYCEPGCBEATVLELASAVKEQPGIETESRLG 60			
Best Local Similarity	100.0%; Pred. No. 1.3e-60;			Db	1 MSGEQQTSAVAPPPEVEPGSGRIVVEYCEPGCBEATVLELASAVKEQPGIETESRLG 60			
Matches	115; Conservative 0; Mismatches 0; Indels 0; Gaps 0.			QY	1 MSGEQQTSAVAPPPEVEPGSGRIVVEYCEPGCBEATVLELASAVKEQPGIETESRLG 60			
QY	1 MSGEQQTSAVAPPPEVEPGSGRIVVEYCEPGCBEATVLELASAVKEQPGIETESRLG 60			Db	1 MSGEQQTSAVAPPPEVEPGSGRIVVEYCEPGCBEATVLELASAVKEQPGIETESRLG 60			
Db	1 MSGEQQTSAVAPPPEVEPGSGRIVVEYCEPGCBEATVLELASAVKEQPGIETESRLG 60			QY	61 GTGAFIEIENGQJLVPFKLENGGPPYKDLIEAIRRASNGETLEKINTNSRPCVIL 115			
QY	61 GTGAFIEIENGQJLVPFKLENGGPPYKDLIEAIRRASNGETLEKINTNSRPCVIL 115			Db	61 GTGAFIEIENGQJLVPFKLENGGPPYKDLIEAIRRASNGETLEKINTNSRPCVIL 115			





**KW** PCOS ovarian cyst; dysmenorrhea; endocrine disorder; pterygium; ovary syndrome; infertility; pregnancy disorder; anovulation; dihydrotestosterone; antiinflammatory; gynaecological; reproductive.  
**KW** inflammatory condition; immune disorder; blood disorder; cardiovascular disorder; respiratory disorder; neurological disorder; gastrointestinal disorder; urinary system disorder; drug screening; gene therapy; chromosome mapping; forensic analysis; antibody preparation; cytostatic; immunomodulatory; neuroprotective; antiinflammatory; cytostatic; immunomodulatory; neuroprotective.  
**KW** antiinflammatory; gynaecological; reproductive.  
**OS** Homo sapiens.  
**XX**  
**PN** WO200200677-A1.  
**XX**  
**PD** 03-JUN-2002.  
**XX**  
**PP** 07-JUN-2001; 2001WO-US018569.  
**PR** 07-JUN-2000; 2000US-0209467P.  
**XX**  
**PA** (HUMAN) HUMAN GENOME SCI INC.  
**XX**  
**PI** Birse CE, Rosen CA;  
**XX**  
**DR** WPI; 2002-147878/19.  
**XX**  
**PT** N-PDB; ABQ56132.  
**XX**  
**PT** Isolated nucleic acid molecules encoding novel ovarian polypeptides, useful in the prevention, treatment and diagnosis of cancer (e.g. ovarian cancer), immune disorders, cardiovascular disorders and neurological diseases.  
**XX**  
**PS** Claim 11; SEQ ID NO 4187; 2922pp; English.  
**CC** The invention relates to 2175 novel human ovarian antigens (ABP41054-ABP4328) and to cDNAs encoding them (ABQ54131-ABQ56305), and also encompasses polypeptides 90% identical and polynucleotides 95% identical to the sequences of the invention. The invention additionally relates to recombinant vectors and host cells comprising human ovarian antigen polynucleotides, antibodies against human ovarian antigens, and the use of ovarian antigen polynucleotides and polypeptides in diagnosing, treating, prognosis or preventing various ovary and/or breast-related disorders. Such conditions include, ovarian cancer and breast cancer, and metastatic tumours of ovarian or breast origin, reproductive system disorders (e.g., infertility, disorders of pregnancy, anovulation, polycystic ovary syndrome, ovarian cysts, and dysmenorrhea), endocrine disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and vaginitis), immune disorders (e.g., congenital and acquired immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus), blood-related disorders (e.g., anaemia), cardiovascular disorders, respiratory disorders, neurological disorders, gastrointestinal disorders and urinary system disorders. Ovarian antigen polypeptides and polynucleotides may also be used in screening for compounds which modulate ovarian antigen expression or activity. The polynucleotides may further be used for gene therapy, chromosome mapping, in the identification of individuals and in forensic analysis, and the polypeptides may be used as food additives or to prepare antibodies useful in disease diagnosis, drug targeting and phenotyping. The present sequence represents a human ovarian antigen of the invention. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at [http://wipo.int/pub/published\\_pct\\_sequences](http://wipo.int/pub/published_pct_sequences)  
**XX**  
**SQ** Sequence 124 AA;  
**OS** Query Match 100.0%; Score 597; DB 5; Length 124;  
**OS** Best Local Similarity 100.0%; Pred. No. 1.4e-60; Length 115; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
**OS** 1 MSGERQQTSAVPPPEVEPGSGVRIWEYCEPCGFEATYLEASAVKEQYQPGIEESLG 60  
**OS** 10 MSGERQQTSAVPPPEVEPGSGVRIWEYCEPCGFEATYLEASAVKEQYQPGIEESLG 69

Qy	61 GTGAFEIENGQVLFSKLENGGPFYEKDLIEARRASNGELEKITSNSRPCVIL 115	Best Local Similarity 100.0%; Pred. No. 1.5e-60; Matches 115; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db	70 GTGAFEIENGQVLFSKLENGGPFYEKDLIEARRASNGELEKITSNSRPCVIL 124	
RESULT 12		
ID	AAB43521	
ID	AAB43521 standard; protein; 131 AA.	
AC	AAB43521;	
XX		
DT	08-FEB-2001 (first entry)	
XX		
DE	Human cancer associated protein sequence SEQ ID NO:966.	
XX		
KW	Human; cancer associated gene; cancer antigen; detection; cancer; diagnosis; cytostatic; proliferative; vasoconstrictive; immunomodulator; antidiabetic; antiinflammatory; antitumour; antiangiogenic; antiviral; antiinflammatory; antithyroid; antiallergic; antibacterial; cardiotonic; dermatological; neuroprotective; thrombolytic; coagulant; nootropic; immune disorder; haemopoietic cell disorder; autoimmune disorder; allergic reaction; graft versus host disease; organ rejection; haemostatic; thrombotic; cardiovascular disorder; infection; neurological disease; drug screening.	
OS	Homo sapiens.	
XX		
PN	WO200055350-A1.	
PD	21-SEP-2000.	
PP	08-MAR-2000; 2000WO-US005882.	
XX		
PR	12-MAR-1999; 99US-0124270P.	
XX		
PA	(HUMA-) HUMAN GENOME SCI INC.	
PT	Rosen CA, Ruben SM;	
XX		
DR	WPI; 2000-587533/55.	
XX		
PT	Novel isolated nucleic acids comprising sequences encoding peptides useful for treating or diagnosing e.g. cancer.	
XX		
PS	Claim 11; Page 1534-1535; 2352pp; English.	
XX		
CC	ABC77607 to AAC7844 encode the human cancer associated proteins given in AAB43398 to AAB4239. The proteins can have activities based on the tissues and cells the genes are expressed in. Example of activities include: cytostatic; proliferative; vasoconstrictive; immunomodulator; antidiabetic; antiasthmatic; antirheumatic; antiarthritic; antiinflammatory; antithyroid; antibacterial; antiviral; dermatological; neuroprotective; cardiotonic; thrombolytic; coagulant; nootropic; vasoconstrictive; antipsoriatic and antiangiogenic. The polypeptides and polypeptides can be used for preventing, treating or ameliorating medical conditions and diagnosing pathological conditions. Polypeptides, polypeptides, antibodies, agonists and antagonists from the present invention may be used to treat immune disorders by activating or inhibiting the proliferation, differentiation or mobilisation of immune cells, to treat disorders of haemopoietic cells, autoimmune disorders, allergic reactions, graft versus host disease and organ rejection, modulate haemostatic or thrombolytic activity, modulate inflammation, cancers, cardiovascular disorders, neurological disease and bacterial or viral infections. The peptides, nucleotides, antibodies, agonists and antagonists may be also be used in drug screens. AAC78449 to AAC78457 and AAC784240 represent sequences used in the exemplification of the present invention	
SQ	Sequence 131 AA;	
Query Match	100.0%; Score 597; DB 8; Length 131;	
	Best Local Similarity 100.0%; Pred. No. 1.5e-60; Matches 115; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
Qy	1 MSGERGQTSVAPPPEVTPSGWRVWVEYRPGFATYFLASAVKEQPGIETSRIG 60	
Db	17 MSGERGQTSVAPPPEVTPSGWRVWVEYRPGFATYFLASAVKEQPGIETSRIG 76	
Qy	61 GTGAFEIENGQVLFSKLENGGPFYEKDLIEARRASNGELEKITSNSRPCVIL 115	
Db	77 GTGAFEIENGQVLFSKLENGGPFYEKDLIEARRASNGELEKITSNSRPCVIL 131	
RESULT 13		
ID	ADK49069	
ID	ADK49069 standard; protein; 131 AA.	
XX		
AC	ADK49069;	
XX		
DT	04-NOV-2004 (first entry)	
XX		
DE	Human breast/bladder carcinoma C35-related protein - SEQ ID 155.	
XX		
KW	C35 epitope; cytostatic; vaccine; tumour; breast; bladder carcinoma; human.	
XX		
OS	Homo sapiens.	
XX		
PN	WO2003104428-A2.	
XX		
PD	18-DEC-2003.	
XX		
PF	10-JUN-2003; 2003WO-US018252.	
XX		
PR	10-JUN-2002; 2002US-0386738P.	
PR	11-DEC-2002; 2002US-0432241P.	
PR	23-APR-2003; 2003US-0464550P.	
XX		
PA	(VACC-) VACCINEX INC.	
PA	(UYRP ) UNIV ROCHESTER.	
XX		
PI	zauderer M, Evans BE, Borrelllo MA;	
XX		
DR	WPI; 2004-062349/06.	
XX		
PT	Novel C35 polypeptide useful for formulation of immunogenic composition to induce antibodies and cell-mediated immunity against tumor cells.	
XX		
PS	Disclosure; SEQ ID NO 155; 62pp; English.	
XX		
CC	The invention relates to a novel isolated polypeptide comprising or consisting of two or more C35 peptide epitopes. The polypeptide of the invention demonstrates cytostatic activity and may be useful for the formulation of an immunogenic composition, such as a vaccine, to induce antibodies and cell-mediated immunity against target cells, such as tumour cells. Furthermore, the polypeptide and its analogues may be useful as prognostic markers for carcinoma, such as human breast or bladder carcinoma. The current sequence is that of human breast/bladder carcinoma C35-related protein of the invention.	
CC	Sequence 131 AA;	
Query Match	100.0%; Score 597; DB 8; Length 131;	
	Best Local Similarity 100.0%; Pred. No. 1.5e-60; Matches 115; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
Qy	1 MSGERGQTSVAPPPEVTPSGWRVWVEYRPGFATYFLASAVKEQPGIETSRIG 60	
Db	17 MSGERGQTSVAPPPEVTPSGWRVWVEYRPGFATYFLASAVKEQPGIETSRIG 76	
Qy	61 GTGAFEIENGQVLFSKLENGGPFYEKDLIEARRASNGELEKITSNSRPCVIL 115	
Db	77 GTGAFEIENGQVLFSKLENGGPFYEKDLIEARRASNGELEKITSNSRPCVIL 131	

RESULT 14

XX AEA62539 standard; protein; 149 AA.

XX ID AEA62539;

XX AC AEA62539;

XX DT 25-AUG-2005 (first entry)

XX QY DR Recombinant C35 antigen.

XX KW C35 antigen; antigen; cell growth; cancer; cytostatic; apoptosis;

XX KW immunotherapy; hyperproliferation; epitope mapping.

XX OS Homo sapiens.

OS Synthetic.

XX

FH KEY Location/Qualifiers

PT Peptide 1. .29 /note= "His-tagged signal peptide"

PT Protein 30. .149 /note= "C35 antigen"

XX PN WO2005055936-A2.

XX

PD 23-JUN-2005.

XX

PP 06-DEC-2004; 2004WO-US040573.

XX

PR 04-DEC-2003; 2003US-0526572P.

PR 23-DEC-2003; 2003US-0531688P.

XX

PA (VACCC-) VACCINEX INC.

XX

PI Evans EE, Paris MJ, Sahasrabudhe DM, Smith ES, Zauderer M;

XX DR WPI; 2005-458501/46.

XX

PT Killing cancer cells, by administering apoptosis-inducing therapy and administering antibody specific for intracellular, cancer-associated protein other than C35, or antibody specific for C35.

XX

PS Example 11; Fig 9; 25pp; English.

XX

CC The invention relates to killing (M1) cancer cells, comprising administering to the cells an antibody specific for an intracellular, cancer-associated protein, provided that the protein is not C35 antigen, where protein becomes exposed on the cell surface in cells undergoing apoptosis, where the antibody is conjugated to or complexed with a toxin. The non-C35 antigen protein is a prenylated protein. Also included are an isolated antibody (1) specific for C35 (chosen from an antibody comprising the VH region encoded by clone 1B3G, the VL region encoded by clone 1B3K, the VH region encoded by clone 1F2G, the VL region encoded by clone 1F2K, the VH region encoded by clone H0009, the VL region encoded by clone L0010, an antibody comprising at least one of CDR1 or CDR2 of the VH region encoded by AB62495, an antibody comprising at least one of CDR1 or CDR2 of the VH region encoded by AB6249, an antibody comprising at least one of CDR1, CDR2, or CDR3 of the VL region encoded by AB62497, a chimeric antibody, or a humanized antibody), a polynucleotide encoding the antibody, a vector comprising the polynucleotide, a host cell comprising the vector and a composition comprising the antibody and a carrier. The method is useful for killing cancer cells in a mammal, preferably human, in need of eradication of smaller tumors and/or micrometastases, or in need of cancer treatment for C35-associated cancer chosen from breast cancer, ovarian cancer, bladder cancer, lung cancer, prostate cancer, pancreatic cancer, colon cancer, melanoma and other hyperproliferative disorders. The antibody is useful for detecting, diagnosing or monitoring C35-associated cancers. The antibody comprises a chimeric antibody comprising human immunoglobulin constant regions fused to the variable regions of mouse anti-35 antibodies (named 1F2, 1B3, Mab 165 and Mab 171). The present sequence represents a recombinant human C35 antigen (with a used His-tagged signal peptide), Lys-C digests of which

CC were used in epitope mapping studies.

XX Sequence 149 AA;

XX

Query Match 100.0%; Score 597; DB 9; Length 149;

XX Best Local Similarity 100.0%; Pred. No. 1; 8e-60; Mismatches 0; Indels 0; Gaps 0;

XX Matches 115; Conservative 115;

QY Db 1 MSGRPQGQSVAPPPEEVPGSGVRIWVXCEPGFEATTELEASAVKQYQPGIEERSLG 60

Db 35 MSGRPQGQSVAPPPEEVPGSGVRIWVXCEPGFEATTELEASAVKQYQPGIEERSLG 94

QY Db 61 GTGAFERIENGQVLFVKLENGGPYKEKOLIEARRASNOETLKINTSRPPCVIL 115

Db 95 GTGAFERIENGQVLFVKLENGGPYKEKOLIEARRASNOETLKINTSRPPCVIL 149

XX

RESULT 15

XX ABR47619 standard; protein; 206 AA.

XX AC ABR47619;

XX DT 12-JUN-2003 (first entry)

XX

DE Breast cancer associated protein sequence SEQ ID NO:480.

XX

KW Human; breast cancer; cytostatic; gene therapy.

XX

OS Homo sapiens.

XX

PN WO2003004989-A2.

XX

PD 16-JAN-2003.

XX

PP 21-JUN-2002; 2002WO-US019669.

XX

PR 21-JUN-2001; 2001US-0299887P.

PR 18-JUL-2001; 2001US-031572P.

PR 28-SEP-2001; 2001US-032500P.

PR 05-MAR-2002; 2002US-0362585P.

PR 14-MAY-2002; 2002US-0380391P.

XX

PA (MILL-). MILLENIUM PHARM INC.

XX

PI Lillie J, Gannavarapu M, Glatt K, Hoersh S, Kamatkar S, Meyers RE, Martens M, Monahan JB, Myer V, Wang Y, Xu Y, Zhao X, Miles GB, Basit RC, Hortobagyi GN, Pusztai L, Meric F, Sahin A, Mills GB;

XX

DR WPI; 2003-210381/20.

XX

NI-PSDB; ACC50321.

XX

PT Breast cancer diagnosis or treatment by comparing the level of expression of a marker in a patient sample with that in the control non-breast cancer sample.

XX

PS Claim 1; SEQ ID NO 480; 128pp; English.

XX

CC The present invention describes a method for assessing whether a patient is afflicted with breast cancer. The method comprises comparing the level of expression of a marker (gene/polypeptide see ACC50076 to ACC50334 and AB4786 to ABR4632) in a patient sample and the normal level of expression of the marker in a control non-breast cancer sample, where a significant increase in the level of expression of the marker in the patient sample and the normal level is an indication that the patient is afflicted with breast cancer. The breast cancer associated sequences from the present invention have cytostatic activities and can be used in gene therapy. The method is useful for diagnosing and treating breast cancer. N.B. The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

XX

SQ sequence 206 AA:

Query Match 100.0%; Score 597; DB 6; Length 206;  
Best Local Similarity 100.0%; Pred. No. 2.8e-60; Mismatches 0; Indels 0; Gaps 0;  
Matches 115; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSGSRGQTSVAPPREVERGSGRIVWEVCEPCGFEATYLELAAVKGQYPPGGBIESRIG 60  
92 MSGSRGQTSVAPPREVERGSGRIVWEVCEPCGFEATYLELAAVKGQYPPGGBIESRIG 151

QY 61 GTRGAFIEIENGQLVFSKUNGGFFYEKDIEARRASNGETLEKITNSRPPCVIL 115  
Db 152 GTRGAFIEIENGQLVFSKUNGGFFYEKDIEARRASNGETLEKITNSRPPCVIL 206

Search completed: March 27, 2006, 12:00:16  
Job time : 121.82 secs





US-11-003-819-2  
; Sequence 2, Application US/11003819  
; Publication No. US20050158323A1  
; GENERAL INFORMATION:  
; APPLICANT: Evans, Elizabeth B.  
; APPLICANT: Paris, Mark J.  
; APPLICANT: sahababudhe, Deepak M.  
; APPLICANT: Zauderer, Maurice  
; APPLICANT: Smith, Ernest S.  
; TITLE OF INVENTION: Methods of Killing Tumor Cells by Targeting Internal Antigens  
; FILE REFERENCE: 1843\_019002  
; CURRENT APPLICATION NUMBER: US/11/003,819  
; CURRENT FILING DATE: 2004-12-06  
; PRIOR APPLICATION NUMBER: US 60/256,572  
; PRIOR FILING DATE: 2003-12-04  
; PRIOR APPLICATION NUMBER: US 60/531,688  
; PRIOR FILING DATE: 2003-12-23  
; NUMBER OF SEQ ID NOS: 61  
; SOFTWARE: PatentIn version 3.3  
; SEQ ID NO: 2  
; LENGTH: 115  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; US-11-003-819-2

RESULT 7  
US-09-833-203-34  
; Sequence 34, Application US/09833203  
; Publication No. US20030116677A1  
; GENERAL INFORMATION:  
; APPLICANT: Zauderer, Maurice  
; APPLICANT: Smith, Ernest S.  
; TITLE OF INVENTION: Targeted Vaccine Delivery Systems  
; FILE REFERENCE: 1821\_002001  
; CURRENT APPLICATION NUMBER: US/09/833,203  
; CURRENT FILING DATE: 2001-04-12  
; PRIOR APPLICATION NUMBER: US 60/196,472  
; PRIOR FILING DATE: 2000-04-12  
; NUMBER OF SEQ ID NOS: 63  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO: 34  
; LENGTH: 117  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; NAME/KEY: misc feature  
; OTHER INFORMATION: C35  
; US-09-833-203-34

Query Match 100.0%; Score 597; DB 6; Length 115;  
Best Local Similarity 100.0%; Pred. No. 2.e-58; Mismatches 115; Conservative 0; Indels 0; Gaps 0;

QY 1 MSGRGQQTSAVAPPREEVEPGSGVRIIVCEPCGFEATYLELASAVKEQYGPGBIESRG 60  
1 MSGRGQQTSAVAPPREEVEPGSGVRIIVCEPCGFEATYLELASAVKEQYGPGBIESRG 60

Db 61 GTGAAFEIENGQLVFSKLENGGFPYEKDIEARRASNGTLEKINTNSRPPCIVL 115  
61 GTGAAFEIENGQLVFSKLENGGFPYEKDIEARRASNGTLEKINTNSRPPCIVL 115

RESULT 8  
US-10-264-049-4187  
; Sequence 4187, Application US/10264049  
; Publication No. US20040005579A1  
; GENERAL INFORMATION:  
; APPLICANT: Birse et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PA133P1  
; CURRENT APPLICATION NUMBER: US/10/264,049  
; CURRENT FILING DATE: 2002-10-04  
; PRIOR APPLICATION NUMBER: PCT/US01/18569  
; PRIOR FILING DATE: 2001-06-07  
; PRIOR APPLICATION NUMBER: US 60/209,467  
; PRIOR FILING DATE: 2000-06-07  
; NUMBER OF SEQ ID NOS: 4360  
; SOFTWARE: PatentIn Ver. 3.1  
; SEQ ID NO: 4187  
; LENGTH: 124  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; US-10-264-049-4187

Query Match 100.0%; Score 597; DB 4; Length 124;  
Best Local Similarity 100.0%; Pred. No. 2.3e-58; Mismatches 0; Conservative 0; Indels 0; Gaps 0;

QY 1 MSGRGQQTSAVAPPREEVEPGSGVRIIVCEPCGFEATYLELASAVKEQYGPGBIESRG 60  
1 MSGRGQQTSAVAPPREEVEPGSGVRIIVCEPCGFEATYLELASAVKEQYGPGBIESRG 60

Db 61 GTGAAFEIENGQLVFSKLENGGFPYEKDIEARRASNGTLEKINTNSRPPCIVL 115  
61 GTGAAFEIENGQLVFSKLENGGFPYEKDIEARRASNGTLEKINTNSRPPCIVL 115

Db 70 GTGAAFEIENGQLVFSKLENGGFPYEKDIEARRASNGTLEKINTNSRPPCIVL 124  
70 GTGAAFEIENGQLVFSKLENGGFPYEKDIEARRASNGTLEKINTNSRPPCIVL 124

RESULT 9  
US-09-925-301-966  
; Sequence 966, Application US/09925301  
; Publication No. US2002005308A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies  
; FILE REFERENCE: PA106  
; CURRENT APPLICATION NUMBER: US/09/925,301  
; CURRENT FILING DATE: 2001-08-10  
; PRIOR APPLICATION NUMBER: PCT/US00/05882  
; PRIOR FILING DATE: 2000-03-08  
; PRIOR APPLICATION NUMBER: 60/124,270  
; PRIOR FILING DATE: 1999-03-12  
; NUMBER OF SEQ ID NOS: 1694  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO: 966  
; LENGTH: 131  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; US-09-925-301-966

Query Match 100.0%; Score 597; DB 3; Length 131;  
Best Local Similarity 100.0%; Pred. No. 2.4e-58; Mismatches 115; Conservative 0; Indels 0; Gaps 0;

QY 1 MSGRGQQTSAVAPPREEVEPGSGVRIIVCEPCGFEATYLELASAVKEQYGPGBIESRG 60  
1 MSGRGQQTSAVAPPREEVEPGSGVRIIVCEPCGFEATYLELASAVKEQYGPGBIESRG 60

Db 61 GTGAAFEIENGQLVFSKLENGGFPYEKDIEARRASNGTLEKINTNSRPPCIVL 115  
61 GTGAAFEIENGQLVFSKLENGGFPYEKDIEARRASNGTLEKINTNSRPPCIVL 115

Db 77 GTGAAFEIENGQLVFSKLENGGFPYEKDIEARRASNGTLEKINTNSRPPCIVL 111  
77 GTGAAFEIENGQLVFSKLENGGFPYEKDIEARRASNGTLEKINTNSRPPCIVL 111

RESULT 11  
US-10-177-293-480  
; Sequence 480, Application US/10177293  
; Publication No. US20030124128A1  
; GENERAL INFORMATION:  
; APPLICANT: Lillie, James  
; APPLICANT: Glatt, Karen  
; APPLICANT: Zhao, Xumei  
; APPLICANT: Ganavararu, Manjula  
; APPLICANT: Kamatkar, Shubhangi  
; APPLICANT: Mertens, Maureen  
; APPLICANT: Myer, Vic  
; APPLICANT: Wang, Youzhen  
; APPLICANT: Xu, Yongyao  
; APPLICANT: Hoersch, Sebastian  
; APPLICANT: Monahan, John  
; APPLICANT: Meyer, Rachel E.  
; APPLICANT: Bast Jr., Robert C.  
; APPLICANT: Hortobagyi, Gabriel N.  
; APPLICANT: Pubzak, LaJos  
; APPLICANT: Meric, Funda  
; APPLICANT: Sahin, Rysegul  
; APPLICANT: Mills, Gordon B.  
; TITLE OF INVENTION: COMPOSITIONS, KITS, AND METHODS FOR IDENTIFICATION, ASSESSMENT, AND THERAPY OF BREAST CANCER  
; FILE REFERENCE: MAI-038  
; CURRENT FILING DATE: 2002-06-21  
; PRIOR APPLICATION NUMBER: US 60/299, 887  
; PRIOR FILING DATE: 2001-06-21  
; PRIOR APPLICATION NUMBER: US 60/301, 572  
; PRIOR FILING DATE: 2001-06-27  
; PRIOR APPLICATION NUMBER: US 60/306, 501  
; PRIOR FILING DATE: 2001-07-18  
; PRIOR APPLICATION NUMBER: US 60/325, 002  
; PRIOR FILING DATE: 2001-09-25  
; PRIOR APPLICATION NUMBER: US 60/362, 585

RESULT 11  
US-10-177-293-480  
; Sequence 480, Application US/10177293  
; Publication No. US20030124128A1  
; GENERAL INFORMATION:  
; APPLICANT: Lillie, James  
; APPLICANT: Glatt, Karen  
; APPLICANT: Zhao, Xumei  
; APPLICANT: Ganavararu, Manjula  
; APPLICANT: Kamatkar, Shubhangi  
; APPLICANT: Mertens, Maureen  
; APPLICANT: Myer, Vic  
; APPLICANT: Wang, Youzhen  
; APPLICANT: Xu, Yongyao  
; APPLICANT: Hoersch, Sebastian  
; APPLICANT: Monahan, John  
; APPLICANT: Meyer, Rachel E.  
; APPLICANT: Bast Jr., Robert C.  
; APPLICANT: Hortobagyi, Gabriel N.  
; APPLICANT: Pubzak, LaJos  
; APPLICANT: Meric, Funda  
; APPLICANT: Sahin, Rysegul  
; APPLICANT: Mills, Gordon B.  
; TITLE OF INVENTION: COMPOSITIONS, KITS, AND METHODS FOR IDENTIFICATION, ASSESSMENT, AND THERAPY OF BREAST CANCER  
; FILE REFERENCE: MAI-038  
; CURRENT FILING DATE: 2002-06-21  
; PRIOR APPLICATION NUMBER: US 60/299, 887  
; PRIOR FILING DATE: 2001-06-21  
; PRIOR APPLICATION NUMBER: US 60/301, 572  
; PRIOR FILING DATE: 2001-06-27  
; PRIOR APPLICATION NUMBER: US 60/306, 501  
; PRIOR FILING DATE: 2001-07-18  
; PRIOR APPLICATION NUMBER: US 60/325, 002  
; PRIOR FILING DATE: 2001-09-25  
; PRIOR APPLICATION NUMBER: US 60/362, 585

RESULT 11  
US-10-177-293-480  
; Sequence 480, Application US/10177293  
; Publication No. US20030124128A1  
; GENERAL INFORMATION:  
; APPLICANT: Lillie, James  
; APPLICANT: Glatt, Karen  
; APPLICANT: Zhao, Xumei  
; APPLICANT: Ganavararu, Manjula  
; APPLICANT: Kamatkar, Shubhangi  
; APPLICANT: Mertens, Maureen  
; APPLICANT: Myer, Vic  
; APPLICANT: Wang, Youzhen  
; APPLICANT: Xu, Yongyao  
; APPLICANT: Hoersch, Sebastian  
; APPLICANT: Monahan, John  
; APPLICANT: Meyer, Rachel E.  
; APPLICANT: Bast Jr., Robert C.  
; APPLICANT: Hortobagyi, Gabriel N.  
; APPLICANT: Pubzak, LaJos  
; APPLICANT: Meric, Funda  
; APPLICANT: Sahin, Rysegul  
; APPLICANT: Mills, Gordon B.  
; TITLE OF INVENTION: COMPOSITIONS, KITS, AND METHODS FOR IDENTIFICATION, ASSESSMENT, AND THERAPY OF BREAST CANCER  
; FILE REFERENCE: MAI-038  
; CURRENT FILING DATE: 2002-06-21  
; PRIOR APPLICATION NUMBER: US 60/299, 887  
; PRIOR FILING DATE: 2001-06-21  
; PRIOR APPLICATION NUMBER: US 60/301, 572  
; PRIOR FILING DATE: 2001-06-27  
; PRIOR APPLICATION NUMBER: US 60/306, 501  
; PRIOR FILING DATE: 2001-07-18  
; PRIOR APPLICATION NUMBER: US 60/325, 002  
; PRIOR FILING DATE: 2001-09-25  
; PRIOR APPLICATION NUMBER: US 60/362, 585

RESULT 12  
US-10-177-293-480  
; Sequence 153, Application US/10457829  
; Publication No. US20040063907A1  
; GENERAL INFORMATION:  
; APPLICANT: Zauderer, Maurice  
; APPLICANT: Evans, Elizabeth E.  
; APPLICANT: Borrello, Melinda A.  
; TITLE OF INVENTION: A Gene Differentially Expressed in Breast and  
; File Reference: 1821-040005  
; Current Application Number: US/10/457, 829  
; Prior Application Number: US 60/454, 650  
; Prior Filing Date: 2003-04-23  
; Number of Seq ID Nos: 160  
; Software: PatentIn Ver. 2.1  
; Seq ID No: 480  
; Length: 206  
; Type: PRT  
; Organism: Homo sapiens  
; US-10-177-293-480

RESULT 12  
US-10-177-293-480  
; Sequence 153, Application US/10457829  
; Publication No. US20040063907A1  
; GENERAL INFORMATION:  
; APPLICANT: Zauderer, Maurice  
; APPLICANT: Evans, Elizabeth E.  
; APPLICANT: Borrello, Melinda A.  
; TITLE OF INVENTION: A Gene Differentially Expressed in Breast and  
; File Reference: 1821-040005  
; Current Application Number: US/10/457, 829  
; Prior Application Number: US 60/454, 650  
; Prior Filing Date: 2003-04-23  
; Number of Seq ID Nos: 160  
; Software: PatentIn Ver. 2.1  
; Seq ID No: 480  
; Length: 206  
; Type: PRT  
; Organism: Homo sapiens  
; US-10-177-293-480

RESULT 13  
US-10-177-293-480  
; Sequence 153, Application US/10457829  
; Publication No. US20040063907A1  
; GENERAL INFORMATION:  
; APPLICANT: Zauderer, Maurice  
; APPLICANT: Evans, Elizabeth E.  
; APPLICANT: Borrello, Melinda A.  
; TITLE OF INVENTION: A Gene Differentially Expressed in Breast and  
; File Reference: 1821-040005  
; Current Application Number: US/10/457, 829  
; Prior Application Number: US 60/454, 650  
; Prior Filing Date: 2003-04-23  
; Number of Seq ID Nos: 160  
; Software: PatentIn Ver. 2.1  
; Seq ID No: 480  
; Length: 206  
; Type: PRT  
; Organism: Homo sapiens  
; US-10-177-293-480

Best Local Similarity 22.0%; Pred. No. 9.7e-05;  
 Matches 38; Conservative 20; Mismatches 27; Indels 88; Gaps 4;

Qy 13 PPEB---VERPGSVRIVYVERCPGFBATYELASAVKEQYQIEI----- 55  
 Db 78 PPEKTTNNATPGSGNTVWIFCSCSYKGATVNMKRMIAFQTEVILANYPPTLPR 137

Qy 56 ----- ESR 58

Db 138 LLSKVUPVWQIGVIGVWVAGEHTPMLGVAPPWVYNNRNGTIASTWLGVALQS 197

Qy 59 LGGTCAFEIBINGOLVFSKLENGGPPYR---KDLIEAIRRASNGETLEKINTS 108

Db 198 LQSSGAFEVYFNGELVFSKLEGRGPPGEIELKDLI-----TKQINTNS 239

RESULT 15  
 US-10-739-930-10642  
 Sequence 10642, Application US/10739930  
 Publication No. US20040216190A1

GENERAL INFORMATION:

APPLICANT: Kovacic, David K.  
 TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH PLANTS AND USES THEREOF FOR PLANT IMPROVEMENT

FILE REFERENCE: 38-21(53377)B  
 CURRENT APPLICATION NUMBER: US/10/739, 930  
 CURRENT FILING DATE: 2003-12-18  
 NUMBER OF SEQ ID NOS: 11088

SEQ ID NO 10642  
 LENGTH: 244

TYPE: PRT  
 ORGANISM: Triticum aestivum  
 FEATURE:  
 OTHER INFORMATION: Clone ID: TRIAB-23APR03-C55262\_1.p

US-10-739-930-10642

Query Match 21.2%; Score 126.5; DB 5; Length 244;  
 Best Local Similarity 21.7%; Pred. No. 1.8e-05; Mismatches 40; Conservative 21; Indels 87; Gaps 5; Matches 40; Mismatches 36; Indels 87; Gaps 5;

Qy 4 EPGQTSVAPPPEEE-----PGGSVRIWYVERCPGFBATYELASAVKEQYQPG 52  
 Db 45 QPGQSS---PDAADSQIAEAQVTDGPSGTTVILKFCJSCSYKGANANTMKRMIDTSFPG 101

Qy 53 IEI----- 55

Db 102 IHVVLENVPPPPKKRAGKIVMPFVQVGAATLMAQDQFPRFRGMVPPPTYSLRANRFGT 161

Qy 55 -----ESLUGGCAFEIBINGOLVFSKLENGGPPYKOLIEA-RRAAN--GE 100  
 Db 162 MASVWMEGNAQSLIQSSGAEVYNGOLVFSKLEQRIPSEFELRELIGNRLSDSQ1GK 221

Qy 101 TLEK 104

Db 222 NLEK 225

RESULT 14  
 US-10-424-599-153780  
 Sequence 153780, Application US/10424599  
 Publication No. US20040031072A1

GENERAL INFORMATION:

APPLICANT: La Rosa Thomas J  
 APPLICANT: Kovacic David K  
 APPLICANT: Zhou Yihua  
 APPLICANT: Cao Yongwei

TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With FILE REFERENCE: 38-21(53223)B  
 CURRENT APPLICATION NUMBER: US/10/424, 599  
 CURRENT FILING DATE: 2003-04-28  
 NUMBER OF SEQ ID NOS: 285684

SEQ ID NO 153780  
 LENGTH: 250

TYPE: PRT  
 ORGANISM: Glycine max  
 FEATURE:  
 OTHER INFORMATION: Clone ID: PAT\_MRT3847\_109887C.1.pep

US-10-424-599-153780

Query Match 21.6%; Score 129; DB 4; Length 250;

Search completed: March 27, 2006, 12:37:11  
 Job time : 100.162 sec